PB-238 555

COLLABORATIVE STUDY OF METHOD FOR THE DETERMINATION OF NITROGEN OXIDE EMISSIONS FROM STATIONARY SOURCES (FOSSIL FUEL FIRED STEAM GENERATORS)

Henry F. Hamil, et al

Southwest Research Institute

Prepared for:

National Environmental Research Center

5 October 1973

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-650/4-74-025	2.
3.	PB 238 555
4. TITLE AND SUBTITLE "Collaborative Study of Method for the Determination of Nitrogen Oxide Emissions from Stationary Sources (Fossil Fuel-Fired Steam Generators)."	5. REPORT DATE October 5, 1973
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Henry F. Hamil and D. E. Camann	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southwest Research Institute 8500 Culebra Rd. San Antonio, Texas 78284	10. PROGRAM ELEMENT NO. 1HA327
	11. CONTRACT/GRANT NO. 68-02-0623
12. SPONSORING AGENCY NAME AND ADDRESS Environmental Protection Agency, QAEML Methods Standardization & Performance Evaluation Branch Research Triangle Park, N. C. 27711	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

A collaborative study was performed on Method 7 proposed by the EPA for determining the nitrogen oxide emissions from stationary sources. Method 7 specifies the collection of a grab sample in an evacuated flask containing a dilute sulfuric acid-hydrogen peroxide absorbing solution, and the colorimetric measurement of the nitrogen oxides, except nitrous oxide, using the phenoldisulfonic acid procedure. Collaborative tests were conducted at both a coal-fired steam generating power plant and an oil-fired pilot plant by four collaborative teams. Statistical analysis of the collaborative test and associated data disclosed the following findings regarding the reliability of a Method 7 performance test result: Precision -- The estimated repeatability standard deviation of a test result is 1.893% of the test result value. The estimated reproducibility standard deviation of a test result is 7.110% of its value. Accuracy -- Because of chemically significant distortions inherent in the gas cylinder accuracy test, the accuracy of Method 7 could not be adequately demonstrated. Minimum Detectable Limit -- The estimated minimum detectable limit of Method 7 is 5.14×10^{-7} lb./s.c.f. Sources of Reproducibility Variation -- Nearly all (93%) of the reproducibility variation in a test result is ascribable to laboratory bias, with the other 7% due to repeatability variation. Most of the apparent laboratory bias variation actually is not a true laboratory effect but rather a day effect primarily caused by dubious daily spectrophotometer re-calibration procedures. Restriction of the spectrophotometer absorbance calibration range to its more accurate upper region can halve the analytical procedure's percentage error contribution to the reproducibility variation.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
US Department of Commerce
Springfield, VA. 22151

PRICES SUBJECT TO CHANGE

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS (This Report)	21. NO. OF PAGES
Unlimited	Unclassified	102
	20. SECURITY CLASS (This page) Unclassified	22. PRICE 5.25

EPA Form 2220-1 (9-73)

EPA-650/4-74-025

COLLABORATIVE STUDY OF METHOD FOR THE DETERMINATION OF NITROGEN OXIDE EMISSIONS FROM STATIONARY SOURCES (FOSSIL FUEL-FIRED STEAM GENERATORS)

by

H. F. Hamil and D. E. Camann

Southwest Research Institute
8500 Culebra Road
San Antonio, Texas 78284

Contract No. 68-02-0623
ROAP No. 26AAG
Program Element No. 1HA327

EPA Project Officer: M. R. Midgett

Quality Assurance and Environmental Monitoring Laboratory
National Environmental Research Center
Research Triangle Park, North Carolina 27711

Prepared for

OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

October 1974

This report has been reviewed by the Environmental Protection Agency and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ABSTRACT

A collaborative study was performed on Method 7 proposed by the Environmental Protection Agency for determining the nitrogen oxide emissions from stationary sources. Method 7 specifies the collection of a grab sample in an evacuated flash containing a dilute sulfuric acid-hydrogen peroxide absorbing solution, and the colorimetric measurement of the nitrogen oxides, except nitrous oxide, using the phenoldisulfonic acid procedure. Collaborative tests were conducted at both a coal-fired steam generating power plant and an oil-fired pilot plant by four collaborative teams. Statistical analysis of the collaborative test and associated data disclosed the following findings regarding the reliability of a Method 7 performance test result:

Precision--The estimated repeatability standard deviation of a test result is 1.893% of the test result value. The estimated reproducibility standard deviation of a test result is 7.110% of its value.

Accuracy--Because of chemically significant distortions inherent in the gas cylinder accuracy test, the accuracy of Method 7 could not be adequately demonstrated.

Minimum Detectable Limit--The estimated minimum detectable limit of Method 7 is 5.14×10^{-7} lb./s.c.f.

Sources of Reproducibility Variation--Nearly all (93%) of the reproducibility variation in a test result is ascribable to laboratory bias, with the other 7% due to repeatability variation. Most of the apparent laboratory bias variation actually is not a true laboratory effect, but rather a day effect primarily caused by dubious daily spectrophotometer re-calibration procedures. Restriction of the spectrophotometer absorbance calibration range to its more accurate upper region can halve the analytical procedure's percentage error contribution to the reproducibility variation.

Various modifications to Method 7 are recommended which should improve its demonstrated precision.

TABLE OF CONTENTS

		Page
I.	INTRODUCTION	1
II.	COLLABORATIVE TESTING OF METHOD 7	2
	A. Collaborative Test Sites	2
	B. Collaborators	8
	C. Philosophy of Collaborative Testing	12
III.	STATISTICAL DESIGN AND ANALYSIS	14
	A. The Experimental Design	14
	B. The Collaborative Test Data	18
	C. The Accuracy of Method 7	21
	D. The Precision of Method 7	29
	E. Accuracy and Precision of the Analytical Procedure	34
	F. The Sources of Variability in Method 7	37
IV.	CONCLUSIONS AND RECOMMENDATIONS	45
	References	51
APPENDICES		
	A. Method 7 - Determination of Nitrogen Oxide Emissions from Stationary Sources	A-1
	B. Statistical Methods	B-1
	B.1 Outlier and Associated Preliminary Analysis of the Original Collaborative Test Data	B-1
	B.2 Significance of the Port Effect	B-6
	B.3 Empirical Relationship of the Standard Deviation to the Mean in the Collaborative Test Data	B-7
	B.4 Transformation of the Collaborative Test Data	B-12
	B.5 The Underlying Relationship of the Standard Deviation Components to the Mean	B-16

Appendices (continued)

		Page
B.6	Estimating the Standard Deviation Components and the Reproducibility Standard Deviation	B-18
B.7	The Nitrate Solution Data	B-22
B.8	The Variance Components of the Analytical Procedure	B-26
B.9	The Lower Limit of Detectability	B-32
	References	B-34
C.	Walden Theoretical NO_x Concentration Calculation	C-1

LIST OF TABLES AND FIGURES

Figure		Page
1.	Pilot Plant Operational Configuration	3
2.	Top View of Test Section	5
3.	Dayton Power and Light Company's Tait Station	6
4.	Test Facilities	7
5.	Stack Gas Delivery System and Sampling Manifold	9
6.	Cambridge NO_x Collaborative Test	10
7.	Dayton NO_x Collaborative Test	11
8.	Collaborative Test of Method 7 - Instructions for Analysis of Unknown Nitrate Solutions	19
9.	The Accuracy of Method 7	28
10.	Analytical Standard Deviation Components $\hat{\sigma}$ and $\hat{\sigma}_L$	38
B1.	Calibration Curve Discrepancies	B-5
B2.	Inter-Laboratory Run Plot	B-10
B3.	Intra-Laboratory Collaborator Block Plot	B-13

Table		
1.	Randomized Block Design of the Cambridge Collaborative Test of Method 7	16
2.	The Original NO_x Collaborative Test Data	20
3.	The Corrected Cambridge Collaborative Test Data	22

LIST OF TABLES AND FIGURES (Cont'd)

Table		Page
4.	The Corrected Dayton Collaborative Test Data	23
5.	Gas Cylinder Test Data Summary	25
6.	Method 7 Accuracy Data	27
7.	Accuracy of the Method 7 Analytical Procedure	35
8.	Sources of Reproducibility Variation in a Performance Test Result	40
9.	The Analytical Procedure as a Source of Component Variation	41
10.	Outlier Collaborative Data Points Derived Outside the Acceptable Calibration Range	43
B1.	Inter-Laboratory Run Summary	B-8
B2.	Intra-Laboratory Collaborator Block Summary	B-11
B3.	Data Transformations to Achieve Run Equality of Variance	B-12
B4.	Reported Nitrate Solution Concentrations, $\mu g\ NO_2$ per 10 ml	B-23
B5.	Day Averaged Nitrate Solution Concentrations, $\mu g\ NO_2$ per 10 ml	B-24
B6.	Average Laboratory Nitrate Solution Concentration, $\mu g\ NO_2$ per 10 ml	B-25
B7.	Nitrate Solution Data Anlayses of Variance	B-27
B8.	Significance of Nitrate Solution Factors	B-28

LIST OF TABLES AND FIGURES (Cont'd)

Table		Page
B9.	Nitrate Solution Variance and Standard Deviation Components	B-31
C1.	Oxygen Consumption for Oil Combustion	C-2
C2.	Component of Flow Due to Stoichiometric Combustion	C-3
C3.	Walden Pilot Plant Firing Conditions	C-3
C4.	Theoretical Calculated NO_x Concentration Results	C-4

I. INTRODUCTION

This report describes the work performed and results obtained on Southwest Research Institute Project 01-3487-001, Contract No. 68-02-0623, which includes collaborative testing of Method 7 for nitrogen oxide emissions as given in "Standards of Performance for New Stationary Sources."(1)

This report describes the collaborative testing of Method 7 in a coal-fired steam generating power plant and in an oil-fired pilot plant, the statistical analysis of the data from the collaborative tests, and the conclusions and recommendations based on the analysis of data.

II. COLLABORATIVE TESTING OF METHOD 7

A. Collaborative Test Sites

Two collaborative tests of Method 7 were conducted. One test was performed at Walden Research Corporation, Cambridge, Massachusetts from December 11 to December 15, 1972, while the second was performed at the Dayton Power and Light Company's Tait Station, Dayton, Ohio from January 8 to January 12, 1973.

The initial test was conducted on the Walden pilot plant since precise control of furnace firing conditions plus accurate addition of nitrogen oxides to the furnace exhaust gas would allow evaluation of the method under carefully controlled nitrogen oxide emission levels. A schematic of the Walden combustion pilot plant is shown in Figure 1. The unit consists of a 400,000 Btu/hour (Jackson and Church) furnace with a combination gas/oil burner. The waste heat is discharged and the exhaust gas from the burner is passed into a series of carbon-steel test sections three feet in length and eight inches in diameter. The flue gas is cooled down to about 300°F by an air-cooled heat exchanger and passed into a second series (3) of carbon-steel test sections (sampling areas). The gas is pulled out of these test sections by a Westinghouse induced-draft fan and exhausted through corrugated pipe at roof level.

The gas doping system consisted of a 1A gas cylinder containing pure nitric oxide, glass rotameter (Fischer & Porter 448-209), and simple

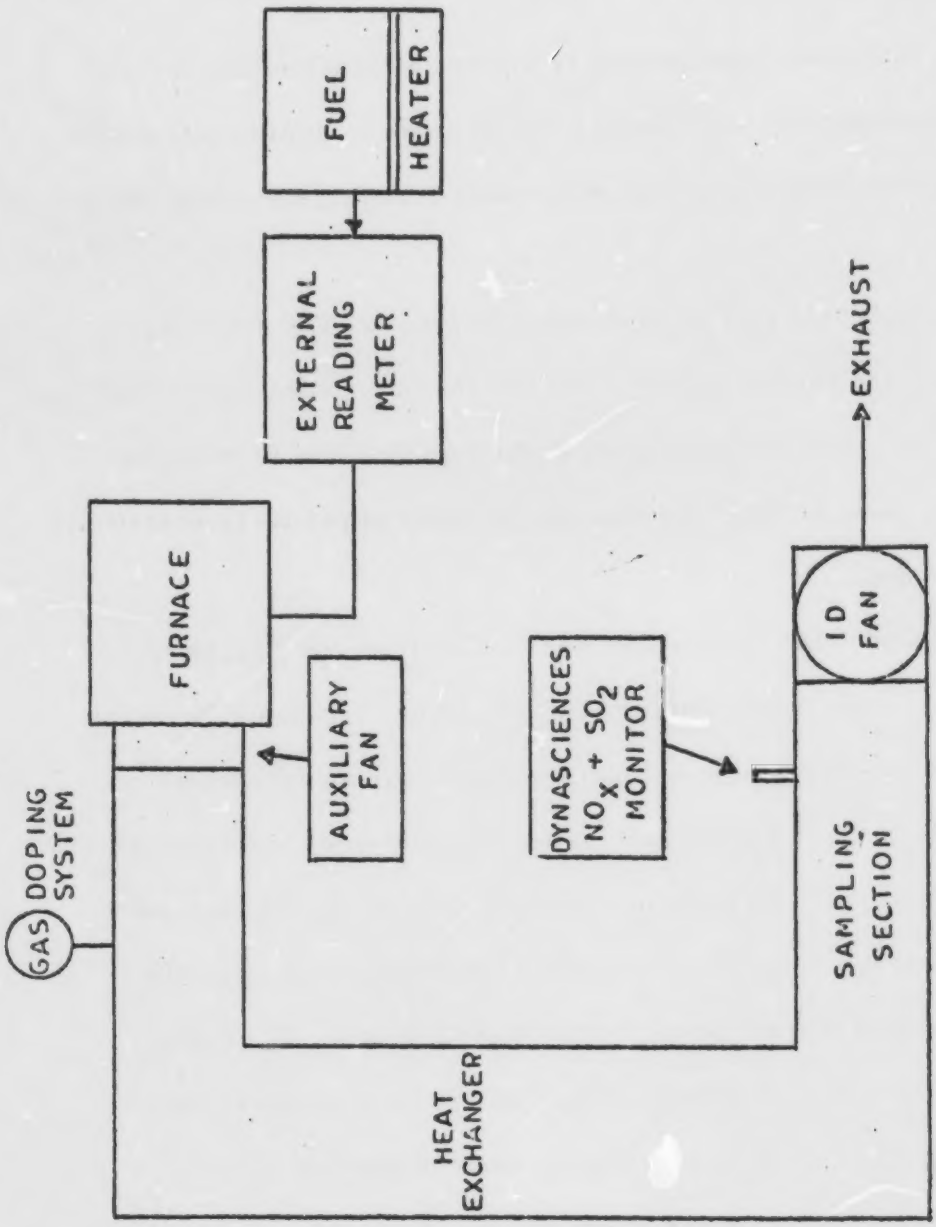


Figure 1. Pilot Plant Operational Configuration.

toggle valve. The dopant gas stream is introduced into the high temperature section immediately after the fire box to come to equilibrium temperature and concentration across the duct before reaching the sample test section.

The sample test section is shown in Figure 2. The sampling ports show both the earlier number code and the letter code, designated on December 12, 1972. The sampling probe was designed to be at the centroid of the duct, although the sample gas velocity profile is essentially flat across the duct.

The second collaborative test was conducted at the Tait Station of Dayton Power and Light Company, Dayton, Ohio. Monsanto Research Corporation and Dayton Power & Light Company have an agreement permitting MRC to use DP&L's Tait Station (Figure 3) for investigation of various instruments and analytical methods for monitoring stationary combustion sources. A 10-ft x 14-ft utility shed was installed on the roof of DP&L's Tait Station between units 4 and 5 (Figure 4). Units 4 and 5 are both tangentially fired, steam boilers burning pulverized coal. The only difference between the two units is that unit 5 now has a set of mirror-grid electrostatic precipitators in operation in addition to the electrostatic precipitators employed on unit 4. The maximum electrical output of each unit is 140 megawatts. The sample delivery line shown in Figure 4 is used to transfer the stack gas from a position

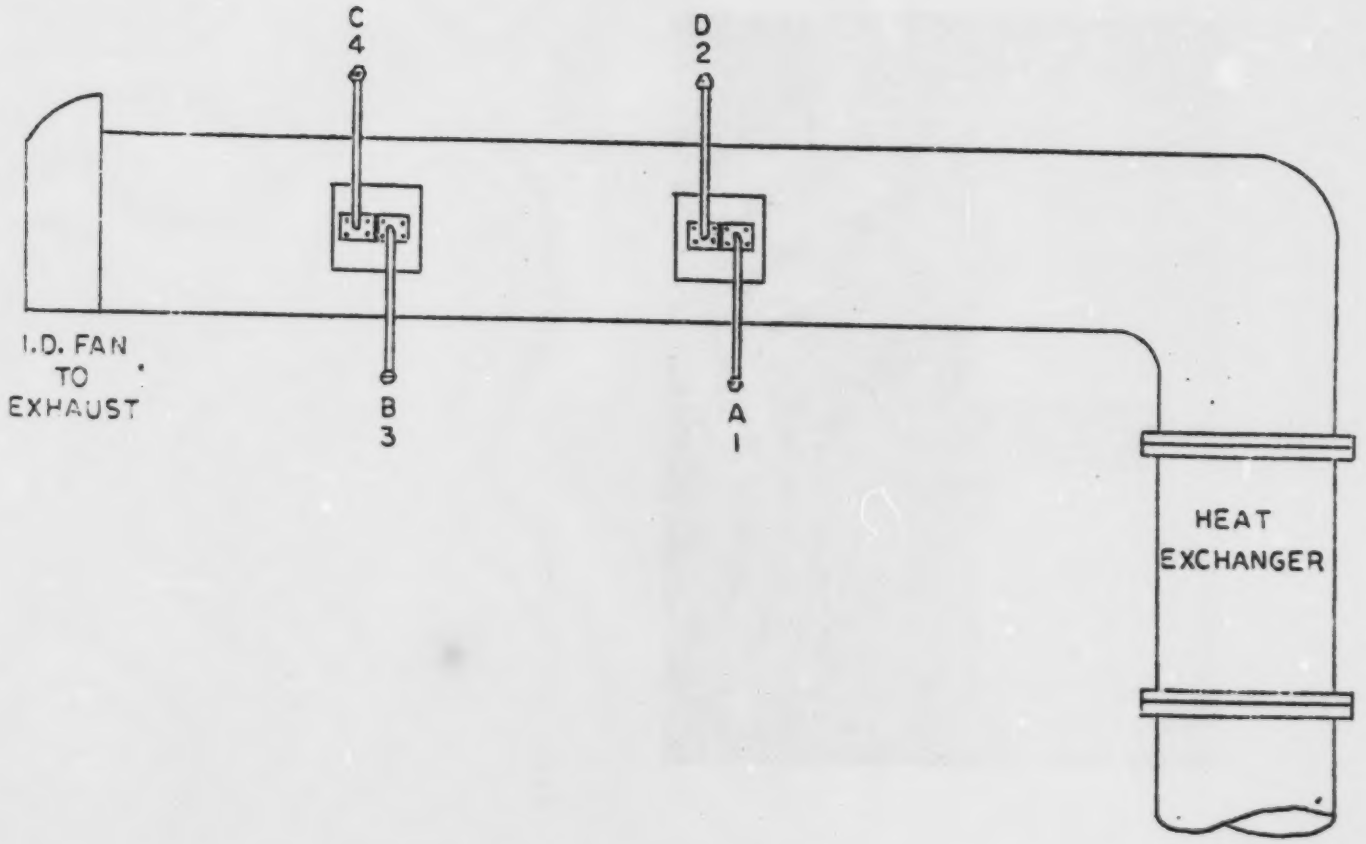


Figure 2. Top View of Test Section.

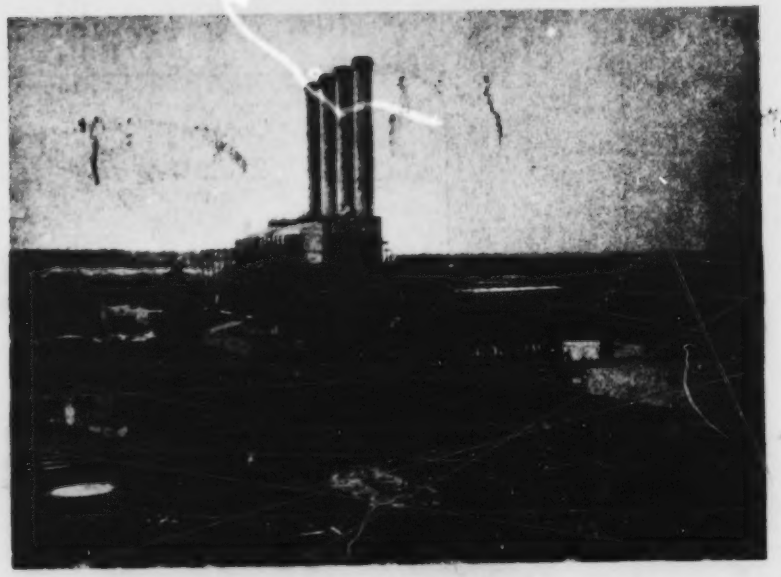


Figure 3. Dayton Power and Light Company's Tait Station

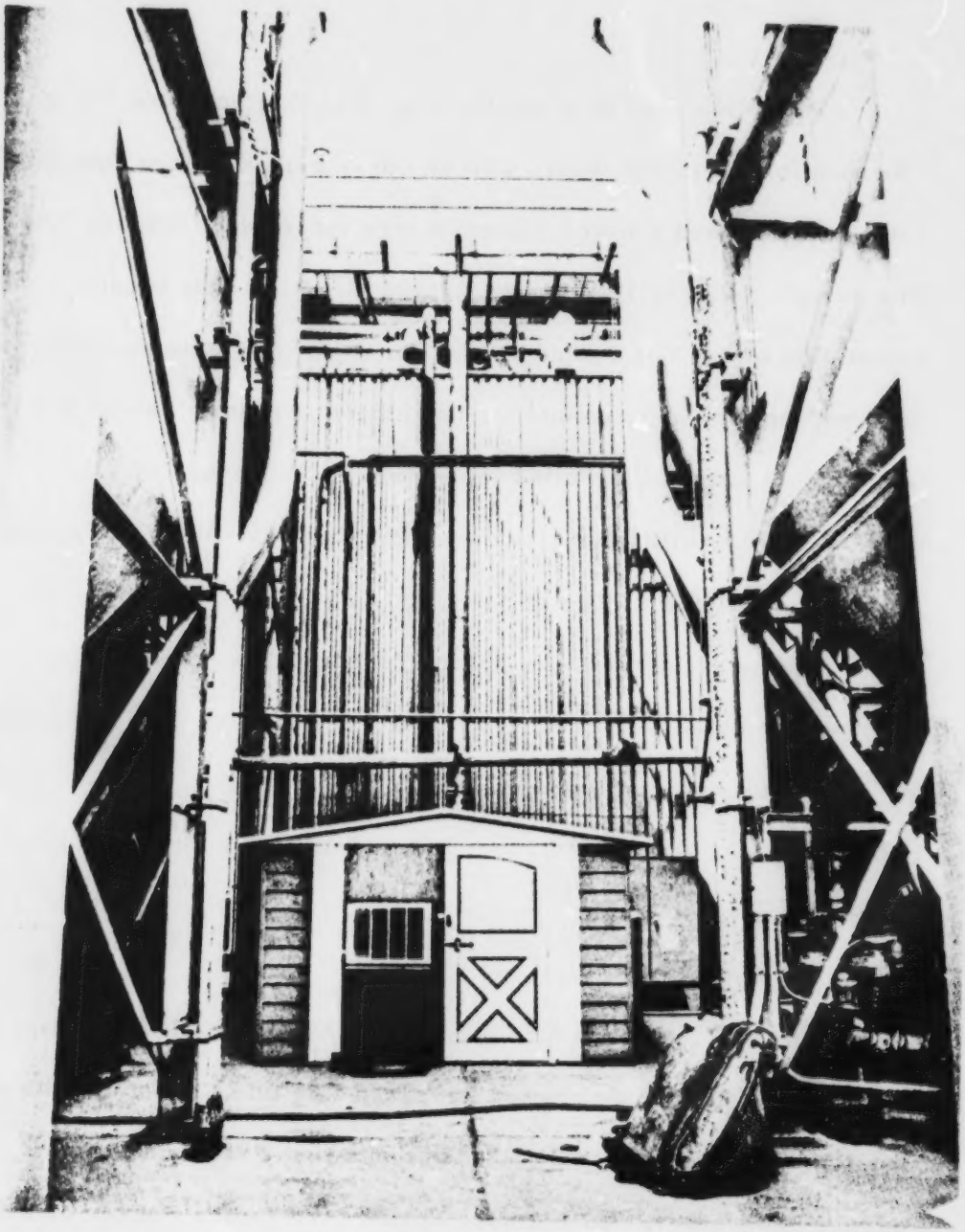


Figure 4. Test Facilities.

after the electrostatic precipitator and before the I. D. fan to the MRC shed.

Inside the shed is a manifold for distribution of the flue gas. The manifold is 10 feet long, with an upper 2-inch square duct fitted with 12 outlets and a lower 8-inch square return duct (Figure 5.). The sample delivery line was connected directly to the manifold for use on this test. The 2-inch black iron connecting pipe was wrapped with heating tape and insulated. The entire system is heated and the temperature can be controlled by sections. Additional sample preparation capabilities include a Rotron Simplex spiral blower for supplying dilution air (Figure 5).

The installation of the Rotron Simplex blower allowed the addition of ambient air to dilute the stack gas to give different levels of nitrogen oxides during the collaborative test. Nitrogen oxide concentration in the stack gas and diluted stack gas was monitored with a calibrated Dynasciences instrument.

B. Collaborators

The collaborators for both the Cambridge and Dayton tests were Mr. Charles Cody of Southwest Research Institute, Houston Laboratory, Houston, Texas; Mr. John Millar of Southwest Research Institute, San Antonio Laboratory, San Antonio, Texas; Mr. James Becker of Walden Research Corporation, Cambridge, Massachusetts; and Mr. Paul Sherman

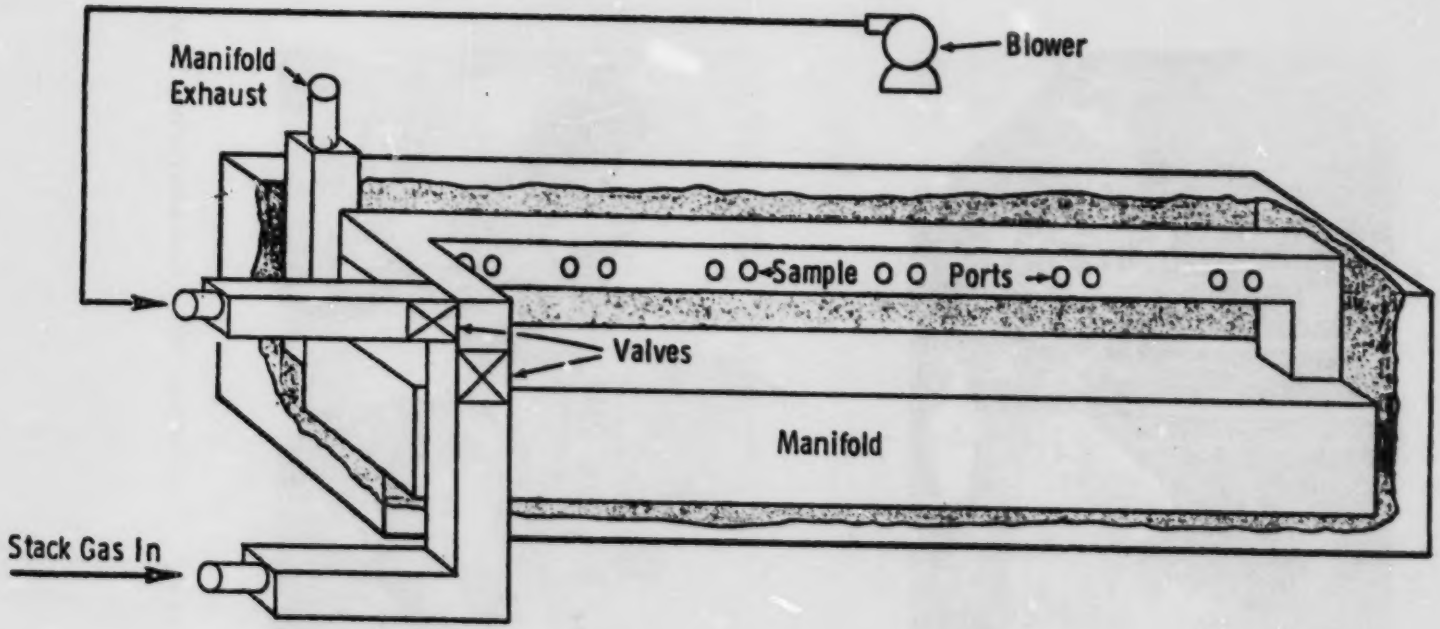
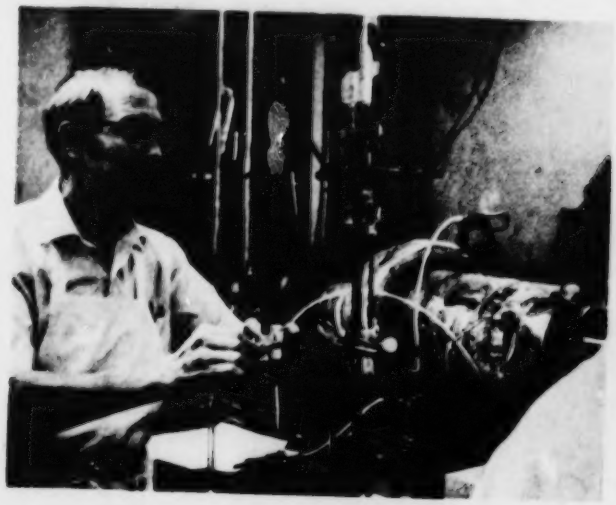


Figure 5. Stack Gas Delivery System and Sampling Manifold.



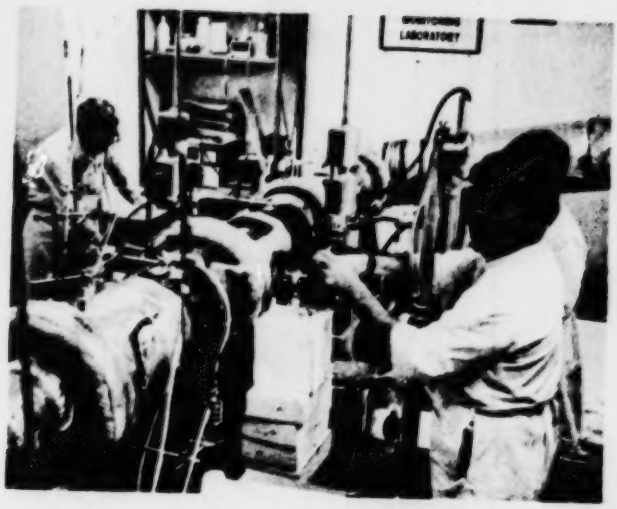


Figure 6. Cambridge NO_x Collaborative Test

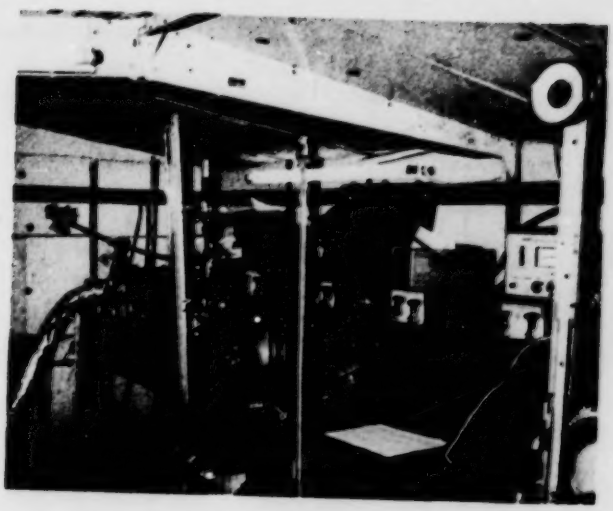


Figure 7. Dayton NO_x Collaborative Test

of Monsanto Research Corporation, Dayton, Ohio. The latter two collaborators were under subcontract to Southwest Research Institute, and in addition to serving as collaborators, had the responsibility for site preparation and test facility maintenance at their respective test sites. Throughout the remainder of this report, the collaborative laboratories are referenced by randomly assigned code numbers as Lab 101, Lab 102, Lab 103, and Lab 104. These code numbers do not correspond to the above ordered listing of collaborators.

Collaborative tests were conducted under the general supervision of Dr. Henry Hamil of Southwest Research Institute. Dr. Hamil had the overall responsibility for assuring that the collaborators were competent to perform the test, that the test was conducted in accordance with the collaborative test plan, and that all collaborators adhered to Method 7 as written in the Federal Register, December 23, 1971.

C. Philosophy of Collaborative Testing

The concept of collaborative testing followed in the tests discussed in this report involves conducting the test in such a manner as to simulate "real world" testing as closely as possible. "Real world" testing implies that the results obtained during the test by each collaborator would be the same results obtainable if he were sampling alone, without outside supervision and without any additional information from outside sources, i.e. test supervisor or other collaborators.

The function of the test supervisor in such a testing scheme is primarily to see that the method is adhered to as written, and that no individual innovations are incorporated into the method by any collaborator. During the test program, the test supervisor observed the collaborators during sampling and sample recovery. If random experimental errors were observed, such as mismeasurement of volume of absorbing solution, improper rinsing of flasks, etc., no interference was made by the test supervisor. Since such random errors will occur in the every day use of this method in the field, unduly restrictive supervision of the collaborative test would bias the method with respect to the performance test results which will be obtained when the method is put into general usage. However, if gross deviations were observed, of such magnitude as to make it clear that the collaborator was not following the method as written, these would be pointed out to the collaborator and corrected by the test supervisor.

While most of the instructions in the Federal Register are quite explicit, some areas are subject to interpretation. Where this was the case, the individual collaborators were allowed to exercise their professional judgement as to the interpretation of the instructions.

The overall basis for this so-called "real-world" concept of collaborative testing is to evaluate the subject method in such a manner as to reflect the reliability, repeatability, and reproducibility of the method that would be expected in performance testing in the field.

III. STATISTICAL DESIGN AND ANALYSIS

A. The Experimental Design

The logistics and other circumstances inherent in emission sampling from the stack of a stationary emissions source impose limitations upon the design of a collaborative test to validate the emission measurement method. In the first place, due to the random nature of stack emissions concentration, dependent as it is upon plant operating level, fuel characteristics, stack flow pattern, etc., the true (i.e.,actual) emission level from the stack is unknown and varies randomly with time despite efforts to maintain the plant in a steady state of operation. As a consequence, there are no true values against which to compare the emissions measurements obtained by the collaborative laboratories performing the method being evaluated. Nor can one assume that the emissions level within the stack remains constant, either in time or in spatial geometry. Because it is physically impossible to locate all four sets of sampling apparatus so as to sample from precisely the same location at the same time, the four sampling teams must sample through different ports in the stack at the same time. While these ports are located so as to be as geometrically symmetrical as possible, the potential for a port effect is nonetheless introduced.

Under these circumstances, the Method 7-Nitrogen Oxide Emissions collaborative test was conducted using a randomized block design at each of the two test sites: Walden Research Corporation's combustion pilot plant in Cambridge, Massachusetts, and Dayton Power & Light Company's

F.M. Tait Station, Dayton, Ohio. A schematic diagram of the randomized block design employed at Cambridge is shown in Table 1. As this table shows, the Cambridge NO_x test was conducted at four different blocks of concentration levels, at approximately 1450 x 10^{-7}, 1000 x 10^{-7}, 675 x 10^{-7}, and 385 x 10^{-7} lb./s.c.f. NO_x concentrations.* Within each block an attempt was made to hold constant the NO_x emission concentration in the stack for the four runs comprising the block. Each run consisted of the simultaneous collection of an NO_x sample from the stack by all four collaborative teams through their assigned port (A, B, C, or D). On the four runs in each block, the collaborative teams rotated from port-to-port in a systematic manner to minimize the logistical problems concomitant with moving the sampling apparatus. Thus in the course of every four run block, each collaborator had sampled once through each port. The four runs in a block were conducted at 15 to 20 minute intervals, just as rapidly as the sampling apparatus could be disconnected, transferred from port to port, and reassembled. This minimized the random ambient variation in the true NO_x emission level in the stack during the collection of a block of data. While the four samples collected by a laboratory team in a block are not replicates in a technical sense (the potential presence of a port effect and the random fluctuation in the true NO_x concentration during the taking of

---

* EPA policy is to express all measurements in Agency documents in metric units. When implementing this practice will result in undue cost or difficulty in clarity, NERC/RTP is providing conversion factors for the particular non-metric units used in the document. For this report, the factor is:

10^{-7} lb/s.c.f. = 1.6017×10^{-3} µg/ml.

TABLE 1. RANDOMIZED BLOCK DESIGN OF THE CAMBRIDGE COLLABORATIVE TEST OF METHOD 7

Block (concentration level) lb/s.c.f.	Run (sample)	Laboratory Team			
		Lab 101	Lab 102	Lab 103	Lab 104
$\sim 1450 \times 10^{-7}$	8	D	B	A	C
	9	A	C	B	D
	10	B	D	C	A
	11	C	A	D	B
$\sim 1000 \times 10^{-7}$	12	A	C	B	D
	13	B	D	C	A
	14	C	A	D	B
	15	D	B	A	C
$\sim 675 \times 10^{-7}$	20	B	D	C	A
	21	C	A	D	B
	22	D	B	A	C
	23	A	C	B	D
$\sim 385 \times 10^{-7}$	30	C	A	D	B
	31	D	B	A	C
	32	A	C	B	D
	33	B	D	C	A

The letters A, B, C, and D depict the sampling ports to which each collaborative laboratory team was assigned on each run. The listed samples were collected for analysis to comprise the sixteen runs of the collaborative test at Cambridge. Samples 1 to 4 were rejected because the glass wool plugs at the sampling ports restricted sample collection. The remaining samples (5-7, 16-19, 24-29) were taken from the standard gas cylinders to constitute the accuracy test.

these samples conflict with the replication concept), nevertheless, within each block the collaborative test has been designed to be closely analogous to the performance test situation for determining compliance via Method 7. Thus, particularly when the port effect can be shown insignificant, these four samples collected by a collaborative team in a block can be considered replicates for the purpose of determining the repeatability of the method. The randomized block design utilized for the Dayton collaborative test was similar to that shown in Table 1., the only important difference being that the four blocks of concentration levels were approximately 465×10^{-7}, 355×10^{-7}, 225×10^{-7}, and 120×10^{-7} lb./s.c.f.

Two ancillary tests were also conducted to study various aspects of the Method 7 collaborative test. One test utilized gas cylinders obtained from Scott Research Laboratories, Inc. containing mixtures of nitric oxide and nitrogen whose nitric oxide concentrations had been accurately determined. The purpose of this gas cylinder test was to provide an independent assessment of the accuracy of Method 7 on sources whose NO_x concentrations, while unknown to the collaborative teams, were reported to the Southwest Research Institute project staff. Three gas cylinders, labeled X, Y, and Z, were used as the NO_x source at both test sites. Along with the four stack samples collected in every block of the collaborative test, each laboratory team also collected samples from each of the three gas cylinders. These samples were later analyzed together with their associated block samples in the laboratory.

A second test involved the replicate analysis of unknown nitrate solutions to determine the reproducibility, repeatability, and minimum detectable limit of the laboratory analytical portion of Method 7. Each of the four laboratories performed triplicate analyses on four different unknown potassium nitrate solutions during three of the days on which the test samples were analyzed. An example of the unknown nitrate solution instruction and reporting sheet is shown in Figure 8.

B. The Collaborative Test Data

The original NO_x collaborative test data reported to the SwRI project staff for the Cambridge test and also for the Dayton test are presented in Table 2. The port from which the sample was collected is shown in parentheses. An outlier analysis was conducted on this data (cf. Appendix B.1) and in the process, numerous detectable calculation errors were uncovered. The data points that are in error are indicated by an asterisk superscript in Table 2. Since calculation errors appear to be so prevalent in obtaining the NO_x concentration using Method 7, and since there is a need to standardize the generation of the absorbance calibration curve (cf. Appendix B1), some thought might be given to the development and use of a standard Method 7 computer program to alleviate these problems. If the Quality Assurance and Environmental Monitoring Laboratory were to design and implement a program to compute the Method 7 NO_x concentration from a laboratory's raw data (a program to include the standardization of measurement units, the generation of a regression calibration line, and

Figure 8.

Collaborative Test of Method 7
Instructions for Analysis of Unknown Nitrate Solutions

A series of nitrate solutions are provided to each collaborator. These solutions are labeled A, B, C, and D, and the concentrations are unknown to the collaborators.

Each unknown solution is to be analyzed in triplicate on each of three separate days. Use a 10 ml aliquot and follow the procedure in Section 5.2 (and 4.3) of Method 7 and report results as micrograms of NO_3 per ml of unknown solution.

Submit the results on this sheet along with your other collaborative test data.

Analyst Lab 104

		Concentration, µg NO_2 per ml			
Day	Replicate	Solution A	Solution B	Solution C	Solution D
Day 1 Date 1-16-73	1	24.8	13.7	39.9	1.0
	2	26.7	12.7	37.0	0.6
	3	26.6	14.1	40.0	0.5
Day 2 Date 1-18-73	1	26.0	12.7	38.6	0.4
	2	24.4	12.9	37.0	0.4
	3	25.8	13.0	38.3	0.4
Day 3 Date 1-22-73	1	25.1	13.0	34.0	0.6
	2	24.4	13.3	34.7	0.6
	3	25.8	12.8	36.2	0.8

TABLE 2. THE ORIGINAL NO_x COLLABORATIVE TEST DATA

Site	Block	Sample	Lab 101 10^{-7} lb/scf	Lab 102 10^{-7} lb/scf	Lab 103 10^{-7} lb/scf	Lab 104 10^{-7} lb/scf
Cambridge						
	1	8	1510*(D)	1443*(B)	1450 (A)	1350 (C)
		9	1520*(A)	1588*(C)	1810 (B)	1350 (D)
		10	1570* (B)	1560*(D)	1500 (C)	1410 (A)
		11	Missing (C)	1653*(A)	1370 (D)	1420 (B)
	2	12	1050*(A)	1486*(C)	1040 (B)	1080 (D)
		13	826* (B)	1107*(D)	1000 (C)	1030 (A)
		14	1010* (C)	1208*(A)	941 (D)	1040 (B)
		15	701* (D)	1166*(B)	996 (A)	960 (C)
	3	20	456*(B)	1240*(D)	763 (C)	670 (A)
		21	554*(C)	715*(A)	750 (D)	660 (B)
		22	696*(D)	727*(B)	806 (A)	670 (C)
		23	734*(A)	798*(C)	697 (B)	700 (D)
	4	30	364*(C)	442*(A)	412 (D)	360 (B)
		31	395*(D)	421*(B)	447 (A)	360 (C)
		32	376*(A)	406*(C)	385 (B)	380 (D)
		33	377*(B)	423*(D)	423 (C)	380 (A)
Dayton						
	1	5	467*(A)	413 (B)	497 (C)	460 (D)
		2	488*(B)	480 (C)	529 (D)	500 (A)
		3	487*(C)	485 (D)	496 (A)	470 (B)
		4	458*(D)	399 (A)	491 (B)	430 (C)
	2	9	335*(B)	321 (C)	362 (D)	350 (A)
		10	339*(C)	296 (D)	373 (A)	360 (B)
		11	365*(D)	356 (A)	421 (B)	390 (C)
		12	359*(A)	348 (B)	408 (C)	380 (D)
	3	16	234*(C)	231 (D)	212 (A)	250 (B)
		17	227*(D)	251 (A)	227 (B)	240 (C)
		18	227*(A)	227 (B)	215 (C)	230 (D)
		19	220*(B)	223 (C)	189 (D)	230 (A)
	4	23	127*(D)	141 (A)	114 (B)	130 (C)
		24	123*(A)	129 (B)	103 (C)	130 (D)
		25	120*(B)	136 (C)	115 (D)	130 (A)
		26	122*(C)	130 (D)	100 (A)	130 (B)

( ) Port from which sample was collected is shown in parentheses.
* Calculation error in this original reported value.

the calculation of both the sample volume under standard conditions and the mass of NO_2), then the calculation error problem could be eliminated. If the program were run by EPA, then the additional problem of bias in performance test reporting (i.e., reporting of only the best twelve of many field test samples in determining compliance) could also be reduced.

The corrected collaborative test data obtained as a result of recalculating suspicious points and correcting the discovered systematic errors are presented in Table 3. (Cambridge data) and Table 4. (Dayton data) along with some initial summary statistics. These statistics are the mean, variance, and standard deviation for each run, collaborative team, and port at each site. The missing Lab 101 point from port C for run 11, block 1 at the Cambridge site has been replaced with the value 1446 in Table 3 obtained as the minimum variance unbiased estimate for Lab 101 on run 11 in block 1 under the variance stabilizing logarithmic transformation (cf. Appendix B.4).

C. The Accuracy of Method 7

A question of vital importance regarding Method 7 is whether this method is accurate in its measurement of the NO_x concentration in the exhaust emitted from the stack of a fossil fuel-fired steam generating power plant, or whether, on the other hand, the method is biased to give consistently low or high measurements. The experiment was designed so

TABLE 3. THE CORRECTED CAMBRIDGE COLLABORATIVE TEST DATA

METHOD: EPA METHOD 7 --- NITROGEN OXIDE EMISSIONS FROM STATIONARY SOURCES

TEST VARIABLE: X = CONCENTRATION OF NOx AS NO2 (DRY BASIS), (LB./S.C.F.)X10**7

TRANSFORMATION: X LINEAR

TEST SITE: CAMBRIDGE

COLLABORATORS: LAB 101 , LAB 102 , LAB 103 , LAB 104 ,

INTER-LABORATORY RUN SUMMARY

RUN	SAMPLE	LAB 101 DATA	LAB 101 PORT	LAB 102 DATA	LAB 102 PORT	LAB 103 DATA	LAB 103 PORT	LAB 104 DATA	LAB 104 PORT	RUN SUMMARY MEAN	RUN SUMMARY STD DEV	RUN SUMMARY BETA
1	8	1440.0	(D)	1337.0	(B)	1450.0	(A)	1350.0	(C)	1394.2	59.0	.0423
2	9	1450.0	(A)	1972.0	(C)	1810.0	(B)	1350.0	(D)	1520.5	200.2	.1316
3	10	1500.0	(B)	1447.0	(D)	1500.0	(C)	1410.0	(A)	1464.2	44.0	.0300
4	11	1446.0	(C)	1531.0	(A)	1370.0	(D)	1420.0	(B)	1441.7	67.3	.0467
5	12	1000.0	(A)	1322.0	(C)	1040.0	(B)	1080.0	(D)	1110.5	144.7	.1303
6	13	789.0	(B)	1027.0	(D)	1000.0	(C)	1030.0	(A)	961.5	115.8	.1204
7	14	963.0	(C)	1120.0	(A)	941.0	(D)	1040.0	(B)	1016.0	81.3	.0800
8	15	669.0	(D)	1081.0	(B)	996.0	(A)	960.0	(C)	926.5	179.0	.1932
9	20	936.0	(B)	736.0	(D)	763.0	(C)	670.0	(A)	651.2	148.7	.2284
10	21	530.0	(C)	663.0	(A)	750.0	(D)	660.0	(B)	650.7	90.7	.1393
11	22	664.0	(D)	675.0	(B)	806.0	(A)	670.0	(C)	703.7	68.3	.0971
12	23	700.0	(A)	790.0	(C)	697.0	(B)	700.0	(D)	709.2	20.5	.0290
13	30	347.0	(C)	911.0	(A)	912.0	(D)	360.0	(B)	382.5	33.9	.0886
14	31	377.0	(D)	341.0	(B)	997.0	(A)	360.0	(C)	393.8	37.7	.0957
15	32	359.0	(A)	374.0	(C)	385.0	(B)	380.0	(D)	374.5	11.3	.0301
16	33	360.0	(B)	392.0	(D)	423.0	(C)	380.0	(A)	388.7	26.9	.0678

COLLABORATOR SUMMARY

COLLABORATOR	LAB 101	LAB 102	LAB 103	LAB 104
MEAN	819.9	919.9	929.9	863.7
STD. DEVIATION	439.1	922.9	929.7	392.8

PORT SUMMARY

PORT	A	B	C	D
MEAN	901.9	873.8	888.7	858.9
STD. DEVIATION	909.3	961.5	925.6	389.5

TABLE 4. THE CORRECTED DAYTON COLLABORATIVE TEST DATA

METHOD: EPA METHOD 7 --- NITROGEN OXIDE EMISSIONS FROM STATIONARY SOURCES

TEST VARIABLE: X = CONCENTRATION OF NOX AS NO2 (DRY BASIS), (LB./S.C.F.)X10**7

TRANSFORMATION: X LINEAR

TEST SITE: DAYTON

COLLABORATORS: LAB 101 , LAB 102 , LAB 103 , LAB 104 ,

INTER-LABORATORY RUN SUMMARY

RUN	SAMPLE	LAB 101 DATA	LAB 101 PORT	LAB 102 DATA	LAB 102 PORT	LAB 103 DATA	LAB 103 PORT	LAB 104 DATA	LAB 104 PORT	RUN SUMMARY MEAN	RUN SUMMARY STD DEV	RUN SUMMARY BETA
1	5	945.0	(A)	913.0	(B)	997.0	(C)	960.0	(D)	953.7	34.9	.0768
2	2	965.0	(B)	980.0	(C)	529.0	(D)	500.0	(A)	993.5	27.7	.0561
3	3	965.0	(C)	985.0	(D)	996.0	(A)	970.0	(B)	979.0	14.2	.0296
4	4	937.0	(D)	999.0	(A)	991.0	(B)	930.0	(C)	939.2	38.2	.0871
5	9	320.0	(B)	321.0	(C)	362.0	(D)	350.0	(A)	338.2	21.1	.0623
6	10	323.0	(C)	296.0	(D)	373.0	(A)	360.0	(B)	338.0	35.1	.1039
7	11	348.0	(D)	356.0	(A)	921.0	(B)	340.0	(C)	378.7	33.5	.0886
8	12	342.0	(A)	348.0	(B)	908.0	(C)	380.0	(D)	369.5	30.6	.0828
9	16	223.0	(C)	231.0	(D)	212.0	(A)	250.0	(B)	229.0	16.0	.0700
10	17	217.0	(D)	251.0	(A)	227.0	(B)	240.0	(C)	233.7	14.9	.0636
11	18	217.0	(A)	227.0	(B)	215.0	(C)	230.0	(D)	222.2	7.4	.0331
12	19	210.0	(B)	223.0	(C)	189.0	(D)	230.0	(A)	213.0	18.0	.0846
13	23	121.0	(D)	141.0	(A)	114.0	(B)	130.0	(C)	126.5	11.7	.0923
14	24	117.0	(A)	129.0	(B)	103.0	(C)	130.0	(D)	119.7	12.6	.1055
15	25	119.0	(B)	136.0	(C)	115.0	(D)	130.0	(A)	123.7	11.0	.0886
16	26	116.0	(C)	130.0	(D)	100.0	(A)	130.0	(B)	119.0	14.3	.1200

COLLABORATOR SUMMARY

COLLABORATOR	LAB 101	LAB 102	LAB 103	LAB 104
MEAN	280.0	285.9	303.2	300.6
STD. DEVIATION	130.3	121.2	154.5	132.5

PORT SUMMARY

PORT	A	B	C	D
MEAN	291.2	293.1	293.8	291.2
STD. DEVIATION	134.2	135.2	138.2	140.3

that the gas cylinder test data would resolve the accuracy question with supporting evidence to be provided by alternative NO_x determination techniques available at the test sites.

Unfortunately, the gas cylinder test proved deficient. A summary of the gas cylinder test data is shown in Table 5. The Method 7 mean is the average of a combined 15 to 20 measurements made by the four collaborative teams on samples taken from each gas cylinder. It is noteworthy that the Method 7 mean is always substantially less than Scott Research's reported true value for the gas cylinder. The remarkable fact is that the percentage difference for the high, medium, and low nitric oxide concentration cylinders is so consistent between the Cambridge and Dayton sites. At both sites, the high NO cylinder was measured 10% low by Method 7, the medium NO cylinder was about 13% low, and the low cylinder was 28% low. As Mr. Paul Sherman of Monsanto Research first suggested, the lack of oxygen in the gas cylinder samples perhaps interfered with the Method 7 determinations. If this were the case, since there is considerable oxygen in real world samples, it would invalidate all the gas cylinder data and consequently negate the gas cylinder approach to determining the accuracy of Method 7.

The lack of molecular oxygen in the cylinder samples does lead to a difference in the total chemistry of the method,[2] as compared to the chemical reactions occuring in oxygen containing samples. In the cylinder samples, the only oxidant present in the flask is the peroxide in the

TABLE 5. GAS CYLINDER TEST DATA SUMMARY

Test Site	Cylinder	NO_x Emission Concentration, 10^{-7} lb./s.c.f. True Value Scott Research, ±1%	Method 7 Mean	Percentage Difference	Number of Measurements
Cambridge					
	High (X)	834	751	-10.0%	16
	Medium (Z)	493	433	-12.2%	20
	Low (Y)	100	72	-28.0%	15
Dayton					
	High (Y)	836	749	-10.4%	16
	Medium (X)	482	415	-13.9%	16
	Low (Z)	96	69	-28.1%	16

absorbing solution. Under these conditions the NO must be absorbed into the liquid phase and subsequently oxidized in stepwise fashion to nitric acid. The overall process is controlled by the rate of diffusion of NO across the gas-liquid interface and the rates of the various oxidation reactions required to convert NO to nitric acid.

In stack samples with oxygen present, some reaction between NO and O_2 will occur, leading to some oxidation in the gas phase in addition to that occurring in the absorbing solution. Because of this, the overall chemistry involved is different for the cylinder samples as compared to the stack samples, and the results obtained for the determination of oxides

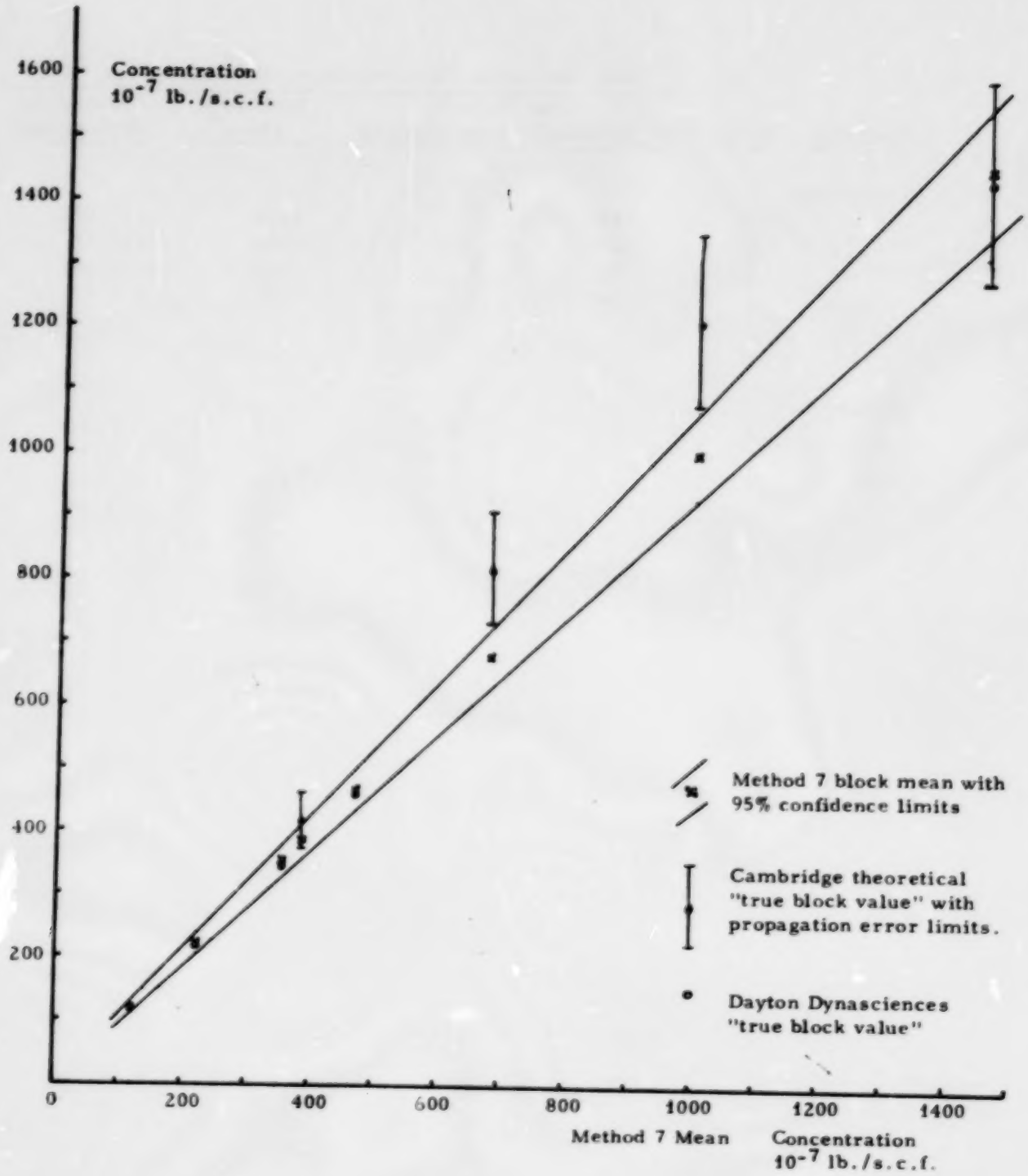
of nitrogen in the presence and absence of oxygen, respectively, are not comparable.

With the gas cylinder test probably invalid, attention must focus on the alternative NO_x determination techniques available and employed at the Cambridge and Dayton sites. Since the Cambridge site was a pilot plant, Walden Research was able to calculate a theoretical concentration of NO_x in the duct at the sample test section based upon the NO doping level, the NO_x due to fuel combustion, and the volumetric flow calculated stoichiometrically. The Walden calculation of theoretical concentration is given in Appendix C. At both the Cambridge and Dayton sites, Dynasciences monitors were available to measure the stack NO_x concentrations. However, the Cambridge Dynasciences monitor was malfunctioning throughout the course of the Cambridge test due to a deficient fuel cell. Thus the Cambridge Dynasciences NO_x readings were not usable. The Cambridge theoretical NO_x concentrations obtained by Walden Research and the Dayton Dynasciences monitor readings obtained by Monsanto are summarized in Table 6 as "true" values for comparison with the corresponding average Method 7 NO_x concentration of the four collaborators in each block. The Table 6 data are plotted in Figure 9. Figure 9 also contains the 95% confidence limits for the Method 7 means of each block and Walden's error range of $\pm$ 11% for the Cambridge theoretical "true block value" obtained by standard propagation of error analysis.(3) The 95% confidence limits for the Method 7 block means were computed using the repeatability variance component σ^2 and the between laboratory variance

TABLE 6. METHOD 7 ACCURACY DATA

Test Site	Block	NO_x Emission Concentration, 10^{-7} lb./s.c.f. "True Value" Theoretical	"True Value" Dynasciences	Method 7 Mean	Percentage Differences
Cambridge					
	1	1440		1455	+ 1.0%
	2	1216		1004	-17.4%
	3	822		679	-17.4%
	4	417		385	- 7.7%
Dayton					
	1		460	466	+ 1.3%
	2		344	356	+ 3.5%
	3		219	224	+ 2.3%
	4		118	122	+ 3.4%

Figure 9. The Accuracy of Method 7



component σ_L^2 estimated in Section III. D. It is evident from Figure 9 that Method 7 and the Dayton Dynasciences monitor were in close agreement throughout the Dayton NO_x test. However, there appears to be a discrepancy between Method 7 and Walden's theoretical calculation method in the Cambridge test, particularly in the second and third blocks at the Method 7 concentrations of 1000×10^{-7} and 680×10^{-7} lb/s.c.f. In considering the Cambridge data, one must view the Walden theoretically calculated "true values" with as much suspicion as the Method 7 means, since neither procedure has been verified. In light of the Method 7—Dynasciences agreement in the Dayton test, it is plausible to suppose that Walden's theoretically calculated "true values" are unreliable. However, some doubt does remain as to the accuracy of Method 7. It is strongly recommended that on future collaborative tests involving Method 7 every reasonable effort be made to ascertain its accuracy.

D. The Precision of Method 7

A major purpose of the NO_x collaborative test is to determine the precision of Method 7. It is usual to express this precision in terms of the repeatability (amount of variability between replicate test results within a single laboratory) and the reproducibility (amount of variability between test results of the same situation obtained in different laboratories) of the method. To maintain consistency with Mandel's definitions of repeatability and reproducibility, [4] the terms repeatability and reproducibility will only be defined in relation to a performance test result for compliance, which is

specified for Method 7 in the Federal Register[5] as the average of three repetitions, each of which consist of four replicate measurements.

Before a detailed statistical analysis of the Method 7 collaborative test data presented in Tables 3 and 4 can be conducted to determine the precision of Method 7, several preliminary matters require settlement. These matters include whether there is a significant port effect to complicate the analysis, whether the between-laboratory run variance and the within-laboratory block variance is mean concentration level dependent, and, if so, what data transformation is required to achieve equality of the run variances.

The Experimental Design section alluded to the potential existence of a significant port effect in the Method 7 tests (i.e., that because of possible non-uniform flow in the stack, the true NO_x emission level at certain ports would be consistently higher or lower than the average of the four ports during the course of the test). The significance of this port effect is tested in Appendix B.2. It is found that no port effect is detectable and consequently, one is justified in ignoring any port effect in the subsequent analysis.

The next problem to be considered is whether the variation in the Method 7 NO_x concentration values is dependent upon the NO_x concentration itself and, if so, what the underlying relationship actually is. It is clearly demonstrated in Appendix B.3 through Figures B.2 and B.3 and through the regression analyses of the data that there is indeed a strong linear relationship

of the standard deviation to the mean spanning both the Cambridge and the Dayton Test data. To be more precise, it appears that the between-laboratory run standard deviation is proportional to the between-laboratory run mean and that the within-laboratory block standard deviation is proportional to the within-laboratory block mean. Appendix B.4 strengthens the argument for this proportional relationship. In Appendix B.4 it is shown that Bartlett's test gives the best equality of between-laboratory run variance by using the logarithmic transformation. And it is further demonstrated that justifying use of the logarithmic transformation on this basis is equivalent to justifying that the between-laboratory run standard deviation is proportional to the between-laboratory run mean. Thus Appendix B.3, with corroborating evidence from Appendix B.4, provides ample justification for concluding that the proportional relationships of standard deviation to mean in both the between laboratory and within laboratory data classifications are underlying characteristics of Method 7 measurements. The consequences of these relationships on the nature of the variance components, σ_L^2 and σ^2, is worked out in Appendix B.5. σ_L^2 is the variance of the population of laboratory means μ_i. σ^2 is the variance of replicate measurements made under similar circumstances by a single laboratory. These definitions correspond to those of Mandel.[6] In Appendix B.5 it is shown that σ and σ_L are both proportional to the mean μ:

$$\sigma = \beta \mu$$

$$\sigma_L = \beta_L \mu$$

with coefficients of variation β and β_L respectively. Therefore, by obtaining estimates of β and β_L from the Cambridge and Dayton collaborative test data, we will have derived estimates of the standard deviation components σ and σ_L through which Mandel defines repeatability and reproducibility.

In Appendix B.6 the methodology for obtaining estimates of β and β_L is derived and applied. The 32 collaborator block point estimates of β given in the last column of Table B2 are averaged and multiplied by the bias correction factor of 1.0854 to yield the within laboratory coefficient of variation estimate $\hat{\beta} = .06558$. So the within lab standard deviation estimate is $\hat{\sigma} = .06558\mu$.

There are 32 run point estimates of the between laboratory coefficient of variation β_b shown in the last column of Table B1. These run point estimates are averaged and corrected by the factor 1.0854 to yield the between laboratory coefficient of variation estimate $\hat{\beta}_b = .09485$. It is shown in Appendix B.6 that $\hat{\beta}_b^2 = \hat{\beta}_L^2 + \hat{\beta}^2$. So $\hat{\beta}_L$ is estimated from $\hat{\beta}_b$ and $\hat{\beta}$ by the equation $\hat{\beta}_L = \sqrt{\hat{\beta}_b^2 - \hat{\beta}^2} = .06853$. Thus the true laboratory-to-laboratory variance component standard deviation estimate is $\hat{\sigma}_L = .06853\mu$.

For a compliance performance test result, which is defined as the average of three repetitions each of four Method 7 measurements, the repeatability standard deviation is $\sigma/\sqrt{m}$ and the reproducibility standard deviation is $\sqrt{\sigma_L^2 + \sigma^2/m}$, where m is the number of measurements which can be considered replicates. When all twelve field samples from the

three repetitions are analyzed in the laboratory on the same day, as currently anticipated, then m = 12 is appropriate. However, if each repetition's four samples are analyzed on different days, then because of the very significant day effect in Method 7 laboratory analysis (cf. Appendix B.8), the proper replication factor is m = 4. For the sake of clarity, the replication factor will hereinafter be assigned the value m = 12. Then the estimated repeatability standard deviation for Method 7 is $\hat{\sigma}/\sqrt{m} = .01893\mu$. With $\phi = 96$ degrees of firmness, the standard deviation of the uncertainty in this repeatability standard deviation estimate is 7.2% of its value. Therefore, the 95% confidence interval for the true repeatability standard deviation $\sigma/\sqrt{m}$ is $.01893\mu \pm (1.96)(.072)(.01893\mu)$ which gives $.01626\mu < \sigma/\sqrt{m} < .02160\mu$. On the basis of the Cambridge and Dayton collaborative tests, there is a 95% chance the actual repeatability standard deviation of a Method 7 test result lies somewhere between 1.63% and 2.16% of its value; the best estimate is 1.89% of this value.

From the Method 7 collaborative data, the estimate of the reproducibility standard deviation that is obtained is $\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/12} = \sqrt{(.06853\mu)^2 + (.06558\mu)^2/12} = .07110\mu$. The estimated reproducibility standard deviation of Method 7 is 7.11% of the test result average. Since it has $\phi = 3.284$ degrees of firmness, the standard deviation of the uncertainty in this reproducibility standard deviation is 39.0%. The 95% confidence interval is $.0168\mu < \sqrt{\sigma_L^2 + \sigma^2/12} < .1254\mu$. Therefore, based upon the Cambridge and Dayton collaborative tests, the probability is 95% that the

Method 7 reproducibility standard deviation lies between 1.68% and 12.54% of the test result average, with a best estimate of 7.11% of the test result average.

Using Mandel's definitions, the estimates of the repeatability and reproducibility of Method 7 test results obtained from the Cambridge and Dayton collaborative tests are as follows:

$$\text{Repeatability} = 2.77\,\hat{\sigma}/\sqrt{m} = 2.77\,(.01893\mu) = .05244\mu$$

$$\text{Reproducibility} = 2.77\,\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/m} = 2.77\,(.07110\mu) = .19694\mu$$

The repeatability of a Method 7 test result is 5.24% of its value, while the reproducibility of a Method 7 test result is 19.69% of its value.

E. Accuracy and Precision of the Analytical Procedure

Four nitrate solutions, whose effective NO_2 source concentrations were known to the project staff but unknown by the collaborative laboratories, were each analyzed in triplicate on each of three days in conjunction with the collaborative test samples by each of the collaborative laboratories. The purpose of this nitrate solution study was to make possible a determination of the accuracy and precision of the laboratory analytical procedure portion of Method 7. All the reported nitrate solution data are presented and summarized in Tables B4, B5, and B6 of Appendix B.7.

TABLE 7. ACCURACY OF THE METHOD 7 ANALYTICAL PROCEDURE

	NO_2 Concentration, µg per 10 ml. of absorbance sample				
Solution	Prepared "True" Value	Collaborator Mean	95% Confidence Interval for Mean	Difference	Percentage Difference
D	0.00	0.83	(0.05, 1.61)	0.83	
B	12.50	13.01	(11.86, 14.16)	0.51	+ 4.08%
A	25.00	25.46	(23.94, 26.98)	0.46	+ 1.84%
C	37.50	38.03	(36.13, 39.93)	0.53	+ 1.41%

The accuracy of the Method 7 analytical procedure is assessed by comparing the average reported NO_2 concentrations of the four solutions for the four collaborative laboratories against the "true" concentration values in terms of the working standard solution of potassium nitrate to which these solutions were prepared. This comparison is presented in Table 7. The 95% confidence interval for the collaborator solution mean is based on a variance of $\hat{\sigma}_L^2/4 + \hat{\sigma}^2/72$ in this mean, where $\hat{\sigma}_L^2$ and $\hat{\sigma}^2$ are the solution's estimated analytical variance components obtained in Appendix B.8. The prepared "true" value for each non-blank solution lies within the 95% confidence interval for its collaborative mean value.

Therefore, the laboratory analytical part of Method 7 appears to be unbiased in the normal working range of the calibration curve. However, the collaborator mean was always somewhat larger than the prepared true value. The difference was actually significant for the blank Solution D. This suggests that the accuracy of the Method 7 analytical procedure deteriorates considerably at very low nitrate concentrations.

A separate analysis of variance was performed on the data for each of the four nitrate solutions to determine the significant factors and to estimate the repeatability variance component. These analyses are reported in Appendix B.8. They show that, for the analytical portion of Method 7, the laboratory-to-laboratory variation is overwhelmingly due, not as might be expected to significant biases from laboratory to laboratory, but rather to large day-to-day variations in measurements that occur within every laboratory. This means that, if the samples collected by the four collaborative teams during the collaborative test had been sent to a single laboratory for analysis on four separate days (instead of to the four respective laboratories for analysis), then the resulting run variation in measurements would be, on the average, nearly as large as that actually obtained. The analytical laboratory variance and standard deviation components estimated for the nitrate solution data are shown in Table B9 of Appendix B.8 for each solution. It is noteworthy that, in contrast to Method 7 itself, the standard deviation components $\hat{\sigma}$ and $\hat{\sigma}_L$ are not proportional to the mean μ. A plot of $\hat{\sigma}$ and $\hat{\sigma}_L$ vs. μ is shown as

Figure 10. It is evident that $\hat{\sigma}_L$, and particularly $\hat{\sigma}$, are linear functions of the mean with a positive intercept at $\mu = 0$. The regression equations are:

$$\hat{\sigma} = .03142\mu + .2321 \text{ µg } NO_2/10 \text{ ml.}$$

$$\hat{\sigma}_L = .0300\mu + .7705 \text{ µg } NO_2/10 \text{ ml.}$$

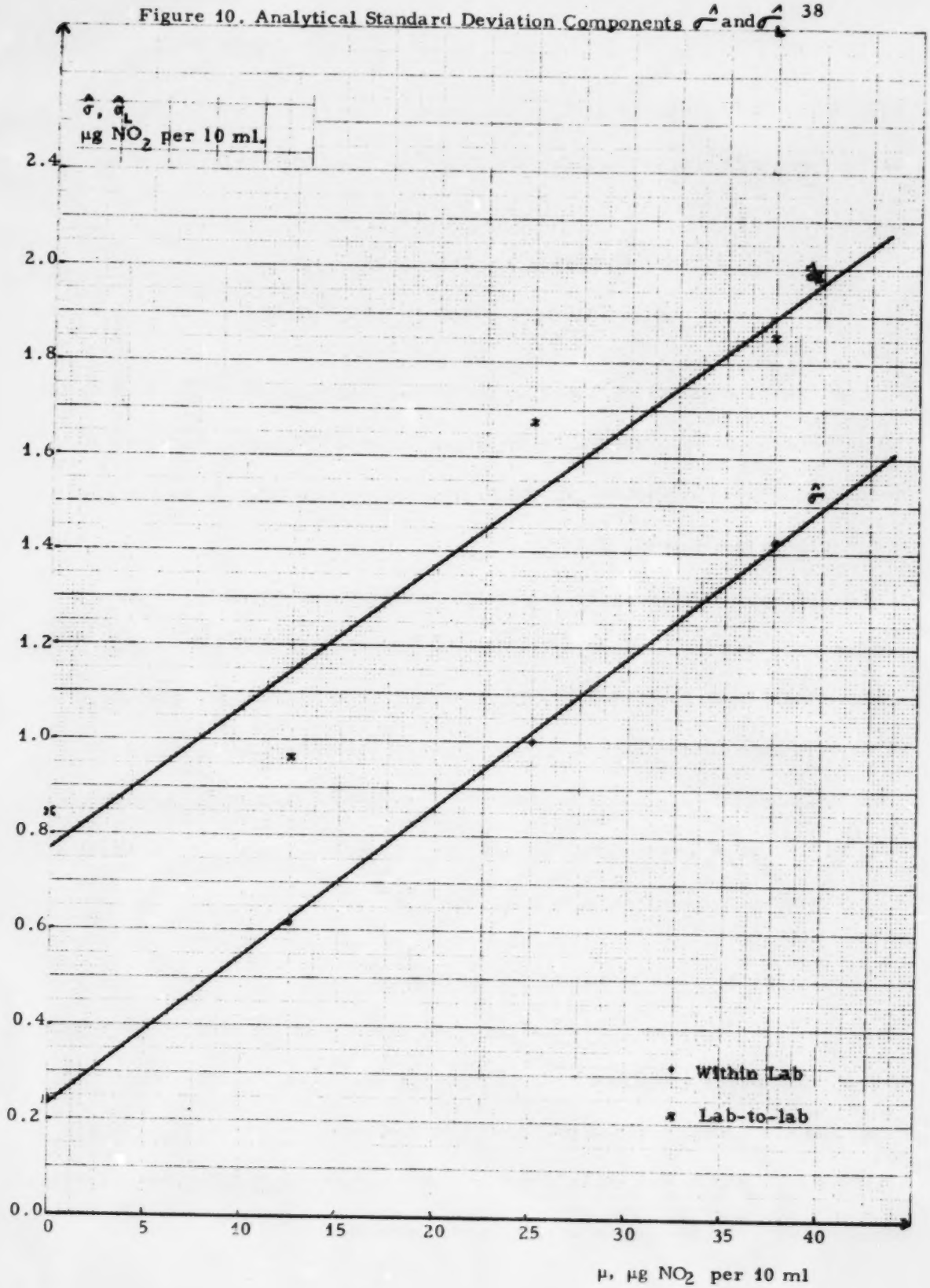
The fact that the standard deviation components for the analytical portion of Method 7 are not proportional to the mean is probably due to the various questionable spectrophotometer calibration curve practices to be discussed more fully in the next section.

The minimum detectable limit for Method 7 is estimated in Appendix B9 as 5.14×10^{-7} lb./s.c.f. This represents the smallest Method 7 performance test result that is significantly larger than a zero nitrogen oxide emission concentration. Thus, if a laboratory obtains a Method 7 performance test result that is less than 5.14×10^{-7} lb/s.c.f. in testing for compliance, such a test result cannot be distinguished from a zero nitrogen oxide emission concentration in the stack.

F. The Sources of Variability in Method 7

The previous statistical analysis has laid the foundation for analyzing the sources of variation in a Method 7 performance test result. The variability in a performance test result is measured by its reproducibility variance $\sigma_L^2 + \sigma^2/12$, which is based on the laboratory-to-laboratory variance component σ_L^2 and the within lab variance component σ^2. Estimates of

Figure 10. Analytical Standard Deviation Components $\hat{\sigma}$ and $\hat{\sigma}_L$



these variance components have been derived, for both Method 7 itself and, from four concentration levels on the spectrophometer calibration curve, for the laboratory analytical portion of Method 7. Since variances of independent factors are additive, it is thus possible to determine to a surprisingly great extent, the Method 7 sources of variability.

The relative importance of σ_L^2 and σ^2 in the variability of a performance test result is shown in Table 8. Almost 93% of the variation in a test result is due to measurement variation from laboratory to laboratory. The repeatability variance component is only 7.1% of the reproducibility variance. The inference is that sufficient replication has been specified in the definition of a Method 7 performance test result.

The percentage of the laboratory-to-laboratory variation, the repeatability variation, and the reproducibility variation attributable to the laboratory analytical and the field portions of Method 7 may be calculated at any point along the absorbance calibration curve. The nitrate solution points of 1.25, 2.50, and 3.75 μgNO_2 per ml. of absorbance sample were used to cover the range of the calibration curve. The results are shown in Table 9. It should be noted that, since the estimates of $\hat{\sigma}_L^2$, $\hat{\sigma}^2/12$, and $\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/12}$ are not very precise, the Table 9 analytical percentages should be considered as rough approximations to reality,

TABLE 8. SOURCES OF REPRODUCIBILITY VARIATION IN A PERFORMANCE TEST RESULT

Variance Component Source	Estimated Variance Component	Variation Percentage
Laboratory-to-laboratory	$\hat{\sigma}_L^2 = (.06853\mu)^2 = .004696\mu^2$	92.9%
Repeatability	$(1/12)\hat{\sigma}^2 = (1/12)(.06558\mu)^2 = .000358\mu^2$	7.1%
Reproducibility	$\hat{\sigma}_L^2 + (1/12)\hat{\sigma}^2 = .005054\mu^2$	100.0%

TABLE 9. THE ANALYTICAL PROCEDURE AS A SOURCE OF COMPONENT VARIATION

	Absorbance Curve Concentration, $\mu gNO_2/ml$.		
	$\mu = 1.25$	$\mu = 2.50$	$\mu = 3.75$
Lab-to-lab: $\hat{\sigma}_L^2$			
Method 7 : $(.06853\mu)^2$	.00734	.02935	.06604
Analytical: $(.03\mu + .07705)^2$	.01312	.02312	.03593
% Analytical	100%	79%	54%
% Field	0%	21%	46%
Repeatability: $\hat{\sigma}^2/12$			
Method 7 : $(.06558\mu)^2/12$	.00056	.00224	.00504
Analytical: $(.03142\mu + .02321)^2/12$	.00033	.00086	.00166
% Analytical	58%	39%	33%
% Field	42%	61%	67%
Reproducibility: $\hat{\sigma}_L^2 + \hat{\sigma}^2/12$			
Method 7	.00790	.03159	.07108
Analytical	.01345	.02398	.03759
% Analytical	100%	76%	53%
% Field	0%	24%	47%

while the field percentages, not being estimated directly, are very unreliable. The main purpose of Table 9 is to indicate that the analytical reproducibility variability can be reduced dramatically by using only the upper portion of the absorbance calibration curve range, from absorbance sample concentrations of, say, 3.0 μg NO_2/ml. to 4.0 μg NO_2/ml. This confirms the assertion developed in Appendix B.1 regarding sample dilution that the low end of the calibration range yields unreliable NO_2 mass data. In fact, many of the outlier collaborative data points that were included in the statistical analysis might, in retrospect, well have been excluded. They might have been excluded on the grounds that they were read from the calibration curve either at the extreme low end of the calibration range where questionable calibration procedures introduced considerable errors, or far above the calibration range where there is no proof that Beers' Law remains valid. These possibly excludable data points are listed in Table 10. If the present Method 7 were revised to require that only the good portion of the calibration curve be used, then the data points listed in Table 10 could be rejected, and the method's repeatability and reproducibility coefficients of variation would be reduced considerably.

The nitrate solution analysis of variance disclosed that nearly all of the analytical laboratory-to-laboratory variance component σ_L^2 is attributable to the day-to-day variation in laboratory measurements. It is likely that the importance of this day-to-day variation derives from

TABLE 9. THE ANALYTICAL PROCEDURE AS A SOURCE OF COMPONENT VARIATION

	Absorbance Curve Concentration, $\mu gNO_2/ml.$		
	$\mu = 1.25$	$\mu = 2.50$	$\mu = 3.75$
Lab-to-lab: $\hat{\sigma}_L^2$			
Method 7 : $(.06853\mu)^2$	.00734	.02935	.06604
Analytical: $(.03\mu + .07705)^2$	.01312	.02312	.03593
% Analytical	100%	79%	54%
% Field	0%	21%	46%
Repeatability: $\hat{\sigma}^2/12$			
Method 7 : $(.06558\mu)^2/12$	.00056	.00224	.00504
Analytical: $(.03142\mu + .02321)^2/12$	.00033	.00086	.00166
% Analytical	58%	39%	33%
% Field	42%	61%	67%
Reproducibility: $\hat{\sigma}_L^2 + \hat{\sigma}^2/12$			
Method 7	.00790	.03159	.07108
Analytical	.01345	.02398	.03759
% Analytical	100%	76%	53%
% Field	0%	24%	47%

TABLE 10. OUTLIER COLLABORATIVE DATA POINTS DERIVED OUTSIDE THE ACCEPTABLE CALIBRATION RANGE

Site	Collaborator	Sample	Concentration lb./s.c.f.	Calibration Curve Point $\mu g\ NO_2$ / ml.
Cambridge	Lab 101	13	789	0.876 - low end of range
Cambridge	Lab 101	15	669	0.766 - low end of range
Cambridge	Lab 102	12	1322	6.64 - above range
Cambridge	Lab 102	15	1081	5.61 - above range
Cambridge	Lab 103	9	1810	5.32 - above range
Dayton	Lab 102	17	251	6.65 - above range
Dayton	Lab 103	19	189	1.14 - low end of range

the daily drift of the spectrophotometer's absorbance readings and the consequent need for daily recalibration. It was noted that the collaborative laboratory's calibration curves were not always revised each day, and that the parameters of the regression line calibration curves were sometimes so imprecise that errors of from 5% to 8% were bound to be introduced into the analytical procedure. It is suggested that these currently allowable but damaging practices be prohibited through a more careful exposition of these matters in the Method 7 procedure.

IV. CONCLUSIONS AND RECOMMENDATIONS

The following conclusions regarding Method 7 (Determination of Nitrogen Oxide Emissions from Stationary Sources) have emerged from the preceding statistical analysis of the Method 7 collaborative test data and associated ancillary data obtained from the Cambridge and Dayton test sites.

1. Accuracy - Because of unanticipated problems due to the chemistry of sampling from a nitric oxide and nitrogen mixture lacking oxygen, the gas cylinder test for determining the accuracy of Method 7 is probably invalid. The supporting accuracy tests based on alternative NO_x determination techniques gave conflicting results; on balance, they appear to support the hypothesis that Method 7 is accurate. However, the accuracy of Method 7 has not been adequately demonstrated.

2. Precision - The precision measures of a Method 7 performance test result, i.e., reproducibility and repeatability, are proportional to the test result itself. The repeatability standard deviation estimate obtained is $\hat{\sigma}/\sqrt{12} = .01893\mu$, and the reproducibility standard deviation estimate is $\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/12} = .07110\mu$, where μ represents the test result average. In the terminology of percentages, the repeatability standard deviation is 1.893% of the test result average and the reproducibility standard deviation is 7.110% of this average. The 95% confidence limits for the true standard deviations are from 1.63% to 2.16% of the average for

repeatability and from 1.68% to 12.54% of the average for reproducibility. The width of the reproducibility confidence interval is an unavoidable penalty caused by the necessary restriction of these collaborative tests to only four simultaneously sampling collaborative teams. According to Mandel's definitions,[7] the repeatability of Method 7 is 5.24% of its test result average and the reproducibility of Method 7 is 19.69% of its test result average.

3. Minimum Detectable Limit - The estimated minimum detectable limit of Method 7 is 5.14×10^{-7} lb./s.c.f. If a performance test result is less than 5.14×10^{-7} lb./s.c.f., the actual nitrogen oxide emission concentration in the stack is indistinguishable from zero.

4. Sources of Reproducibility Variation - The Method 7 performance test result reproducibility variation can be partitioned into its components and thereby ascribed to its constituent sources. 92.9% of the test result reproducibility variation is accounted for by laboratory bias component σ_L^2, with the remaining 7.1% accounted for by the repeatability component $\sigma^2/12$. With respect to the laboratory analysis phase as opposed to the field sampling phase of Method 7, from 50% to nearly 100% of the reproducibility variation is attributable to the analytical phase, depending upon what portion of the spectrophotometer's absorbance calibration range is used. Use of the upper part of the calibration range minimizes the variability caused by the analytical phase. Finally, almost all of the

analytical portion of the laboratory bias component is not actually a true laboratory effect at all, but rather a day effect mainly attributable to poor or neglected daily spectrophotometer re-calibration procedures.

Assessments of Method 7 have been made by the collaborative test supervisor and by the collaborators themselves as a result of their observations and experience in conducting the field and analytical phases of Method 7. These assessments have included the following pertinent comments.

1. Method 7 is fairly tedious and time-consuming. As a result there are ample opportunities for errors to be introduced, particularly in sample collection and aliquoting.

2. Walden Research Corporation reports that Method 7 for NO_x, as written, is a good test procedure with relatively few weak points. However, one of these is that the line to the manometer should be left open during sampling so that any problems with the glass wool plugging or probe malfunction can be readily observed. The system at this point has already been leak-tested so there is no danger of sample contamination from the manometer side. Secondly, it should be emphasized in the procedure that only a steam bath be used for evaporation of the samples as hot spots generated with a hot plate could introduce error in the determination. Thirdly, it has been their experience that certain types of litmus paper, when dipped into the solutions during the procedure to check alkalinity, introduce an error. The solution should be checked by dipping a glass rod in the

solution and touching it to the litmus paper, wiping the glass rod and repeating the step until the desired results are achieved. This error is greater at lower NO_x levels. Lastly, due to the many handling steps and chance for mishap, it is strongly recommended that an aliquoting section be inserted in the procedure. Aliquoting of samples is a basic procedure in analytical chemistry and would help in the determination of precision in the results. It would also guard against loss of sample and data if a mishap occurs in analysis.

3. Monsanto Research Corporation had two worthwhile suggestions. In the concentration calculations, the constant 6.2×10^{-5} should be expanded to include the second digit after the decimal : $6.24 \times 10^{-5} \frac{lb/scf}{\mu g/ml}$. There is no justification for rounding this number to 6.2 since the number does come out 6.243 when 1.0 is divided by 1.6017. An important point is that when the spectra of the standards and samples generated using this method is scanned on a calibrated dual beam instrument, the maximum absorbance does not come at 420 nm as listed in method. The maximum is at 405 nm. It is, therefore, proposed that the method be changed to read 410 nm or 405 nm. The instruments on which these measurements were made had been calibrated with respect to wavelength using holmium oxide glass and the absorbance was calibrated using a standard absorbance filter provided by the instrument manufacturer. In addition, the 13th Edition of Standard Methods of Analysis of Water and Wastewater-Nitrate Analysis, 1971 and ASTM Standards, Part 23, ASTM designation D-1608-60 recommend 410 nm and 400 nm, respectively. It is always best to do an analysis at the peak maximum instead of on the side of the peak.

The conclusions and pertinent comments presented above, provide a firm basis for the following recommendations concerning Method 7:

1. Additional collaborative evidence is needed to verify the accuracy of Method 7. Consequently, valid accuracy tests should be developed for inclusion in future collaborative tests of Method 7 to provide the necessary evidence.

2. It is recommended that the current version of Method 7 be revised to incorporate these concepts:

a. Conduct a thorough critical review of Method 7 as currently written to locate ambiguous statements and to modify them so as to be more explicit.

b. Calculation errors were prevalent in the Method 7 collaborative test data. To prevent the occurrence of calculation errors in the Method 7 performance test results for compliance, it is recommended that a general Method 7 computer program be written to calculate the NO_x concentration test result from the raw field and laboratory data. All data processing by this Method 7 program should be controlled by the Environmental Protection Agency. In addition to insuring proper program usage, data processing by EPA would also discourage the illegal practice of taking numerous field samples during the performance test and reporting only the twelve best measurements, because the testing laboratory need never calculate the concentration of its samples.

c. Provide more detail regarding the proper spectrophotometer calibration procedure. These details ought to include a requirement for daily re-calibration and generation of the appropriate calibration line. This calibration line should be forced to pass through the origin, either graphically or by linear regression. If a regression line is computed, at least three significant digits should be maintained in the calculated slope.

d. Restrict use of the calibration line to only the more accurate portion of the calibration range. If 2.c above is enacted, the range from 1.0 to 4.0 μg NO_2 per ml. of absorbance sample is considered accurate; otherwise only the range from 2.5 to 4.0 μg NO_2 per ml. should be used. If calibration data are collected between 4.0 and 5.0 μg NO_2 per ml. and the relationship remains linear, then the effective ranges above could be extended from 4.0 to 5.0 μg NO_2 per ml. as the upper limit. This argument is based on the use of absorbance cells with a 1 cm. path length.

Enactment of these four recommendations could greatly enhance both the repeatability, and especially the reproducibility, that are herein reported for the current version of Method 7.

References

1. Environmental Protection Agency, "Standards of Performance for New Stationary Sources," Federal Register, Vol. 36, No. 247, Dec. 23, 1971, pp. 24876-24893.

2. Margolis, Geoffrey, and Driscoll, John N., "Critical Evaluation of Rate-Controlling Processes in Manual Determination of Nitrogen Oxides in Flue Gasas," Environmental Science and Technology, Vol. 6, No. 8, Aug. 1972, pp. 727-731.

3. Strobel, H.A., Chemical Instrumentation, Addison-Wesley, Cambridge, 1960, pp. 27-29.

4. Mandel, John, "Repeatability and Reproducibility," Materials Research and Standards, American Society for Testing and Materials, Vol. II, No. 8, August, 1971, pp. 11, 12.

5. Environmental Protection Agency, op. cit. pp. 24878, 24879.

6. Mandel, John, op. cit., p. 9.

7. Mandel, John, op. cit., p. 12.

APPENDIX A

Method 7. Determination of Nitrogen Oxide Emissions from Stationary Sources

Federal Register, Vol. 36, No. 247
December 23, 1971

METHOD 7—DETERMINATION OF NITROGEN OXIDE EMISSIONS FROM STATIONARY SOURCES

1. *Principle and applicability.*

1.1 Principle. A grab sample is collected in an evacuated flask containing a dilute sulfuric acid-hydrogen peroxide absorbing solution, and the nitrogen oxides, except

nitrous oxide, are measure colorimetrically using the phenoldisulfonic acid (PDS) procedure.

1.2 Applicability. This method is applicable for the measurement of nitrogen oxides from stationary sources only when specified by the test procedures for determining compliance with New Source Performance Standards.

2. *Apparatus.*

2.1 Sampling. See Figure 7-1.

2.1.1 Probe—Pyrex[1] glass, heated, with filter to remove particulate matter. Heating is unnecessary if the probe remains dry during the purging period.

2.1.2 Collection flask—Two-liter, Pyrex,[1] round bottom with short neck and 24/40 standard taper opening, protected against implosion or breakage.

[1] Trade name.

2.1.3 Flask valve—T-bore stopcock connected to a 24/40 standard taper joint.

2.1.4 Temperature gauge—Dial-type thermometer, or equivalent, capable of measuring 2° F. intervals from 25° to 125° F.

2.1.5 Vacuum line—Tubing capable of withstanding a vacuum of 3 inches Hg absolute pressure, with "T" connection and T-bore stopcock, or equivalent.

2.1.6 Pressure gauge—U-tube manometer, 36 inches, with 0.1-inch divisions, or equivalent.

2.1.7 Pump—Capable of producing a vacuum of 3 inches Hg absolute pressure.

2.1.8 Squeeze bulb—One way.

2.2 Sample recovery.

2.2.1 Pipette or dropper.

2.2.2 Glass storage containers—Cushioned for shipping.

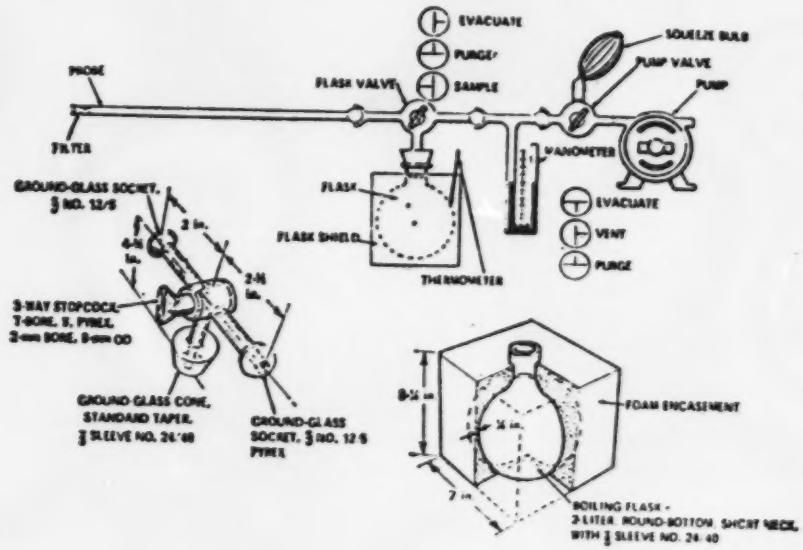


Figure 7-1. Sampling train, flask valve, and flask.

2.2.3 Glass wash bottle.

2.3 Analysis.

2.3.1 Steam bath.

2.3.2 Beakers or casseroles—250 ml., one for each sample and standard (blank).

2.3.3 Volumetric pipettes—1, 2, and 10 ml.

2.3.4 Transfer pipette—10 ml. with 0.1 ml. divisions.

2.3.5 Volumetric flask—100 ml., one for each sample, and 1,000 ml. for the standard (blank).

2.3.6 Spectrophotometer—To measure absorbance at 420 nm.

2.3.7 Graduated cylinder—100 ml. with 1.0 ml. divisions.

2.3.8 Analytical balance—To measure to 0.1 mg.

3. *Reagents.*

3.1 Sampling.

3.1.1 Absorbing solution—Add 2.8 ml. of concentrated H_2SO_4 to 1 liter of distilled water. Mix well and add 6 ml. of 3 percent hydrogen peroxide. Prepare a fresh solution weekly and do not expose to extreme heat or direct sunlight.

3.2 Sample recovery.

3.2.1 Sodium hydroxide (1*N*)—Dissolve 40 g. NaOH in distilled water and dilute to 1 liter.

3.2.2 Red litmus paper.

3.2.3 Water—Deionized, distilled.

3.3 Analysis.

3.3.1 Fuming sulfuric acid—15 to 18% by weight free sulfur trioxide.

3.3.2 Phenol—White solid reagent grade.

3.3.3 Sulfuric acid—Concentrated reagent grade.

3.3.4 Standard solution—Dissolve 0.5495 g. potassium nitrate (KNO_3) in distilled water and dilute to 1 liter. For the working standard solution, dilute 10 ml. of the resulting solution to 100 ml. with distilled water. One ml. of the working standard solution is equivalent to 25 μg. nitrogen dioxide.

3.3.5 Water—Deionized, distilled.

3.3.6 Phenoldisulfonic acid solution—Dissolve 25 g. of pure white phenol in 150 ml. concentrated sulfuric acid on a steam bath. Cool, add 75 ml. fuming sulfuric acid, and heat at 100° C. for 2 hours. Store in a dark, stoppered bottle.

4. *Procedure.*

4.1 Sampling.

4.1.1 Pipette 25 ml. of absorbing solution into a sample flask. Insert the flask valve stopper into the flask with the valve in the "purge" position. Assemble the sampling train as shown in Figure 7-1 and place the probe at the sampling point. Turn the flask valve and the pump valve to their "evacuate" positions. Evacuate the flask to at least 3 inches Hg absolute pressure. Turn the pump valve to its "vent" position and turn off the pump. Check the manometer for any fluctuation in the mercury level. If there is a visible change over the span of one minute, check for leaks. Record the initial volume, temperature, and barometric pressure. Turn the flask valve to its "purge" position, and then do the same with the pump valve. Purge the probe and the vacuum tube using the squeeze bulb. If condensation occurs in the probe and flask valve area, heat the probe and purge until the condensation disappears. Then turn the pump valve to its "vent" position. Turn the flask valve to its "sample" position and allow sample to enter the flask for about 15 seconds. After collecting the sample, turn the flask valve to its "purge" position and disconnect the flask from the sampling train. Shake the flask for 5 minutes.

4.2 Sample recovery.

4.2.1 Let the flask set for a minimum of 16 hours and then shake the contents for 2 minutes. Connect the flask to a mercury filled U-tube manometer, open the valve from the flask to the manometer, and record the flask pressure and temperature along with the barometric pressure. Transfer the flask contents to a container for shipment or to a 250 ml. beaker for analysis. Rinse the flask with two portions of distilled water (approximately 10 ml.) and add rinse water to the sample. For a blank use 25 ml. of absorbing solution and the same volume of distilled water as used in rinsing the flask. Prior to shipping or analysis, add sodium hydroxide (1*N*) dropwise into both the sample and the blank until alkaline to litmus paper (about 25 to 35 drops in each).

4.3 Analysis.

4.3.1 If the sample has been shipped in a container, transfer the contents to a 250 ml. beaker using a small amount of distilled water. Evaporate the solution to dryness on a steam bath and then cool. Add 2 ml. phenoldisulfonic acid solution to the dried residue and triturate thoroughly with a glass rod. Make sure the solution contacts all the residue. Add 1 ml. distilled water and four drops of concentrated sulfuric acid. Heat the solution on a steam bath for 3 minutes with occasional stirring. Cool, add 20 ml. distilled water, mix well by stirring, and add concentrated ammonium hydroxide dropwise with constant stirring until alkaline to litmus paper. Transfer the solution to a 100 ml. volumetric flask and wash the beaker three times with 4 to 5 ml. portions of distilled water. Dilute to the mark and mix thoroughly. If the sample contains solids, transfer a portion of the solution to a clean, dry centrifuge tube, and centrifuge, or filter a portion of the solution. Measure the absorbance of each sample at 420 nm. using the blank solution as a zero. Dilute the sample and the blank with a suitable amount of distilled water if absorbance falls outside the range of calibration.

5. *Calibration.*

5.1 Flask volume. Assemble the flask and flask valve and fill with water to the stopcock. Measure the volume of water to ±10 ml. Number and record the volume on the flask.

5.2 Spectrophotometer. Add 0.0 to 16.0 ml. of standard solution to a series of beakers. To each beaker add 25 ml. of absorbing solution and add sodium hydroxide (1*N*) dropwise until alkaline to litmus paper (about 25 to 35 drops). Follow the analysis procedure of section 4.3 to collect enough data to draw a calibration curve of concentration in μg. NO_2 per sample versus absorbance.

6. *Calculations.*

6.1 Sample volume.

$$V_{sc}=\frac{T_{std}(V_f-V_s)}{P_{std}}\left(\frac{P_f}{T_f}-\frac{P_i}{T_i}\right)=\left(17.71\ \frac{°R}{\text{in. Hg}}\right)(V_f-25\text{ ml.})\left(\frac{P_f}{T_f}-\frac{P_i}{T_i}\right)\quad\text{Equation 7-1}$$

where:

V_{sc} = Sample volume at standard conditions (dry basis), ml.

T_{std} = Absolute temperature at standard conditions, 530° R.

P_{std} = Pressure at standard conditions, 29.92 inches Hg.

V_f = Volume of flask and valve, ml.

V_s = Volume of absorbing solution, 25 ml.

P_f = Final absolute pressure of flask, inches Hg.

P_i = Initial absolute pressure of flask, inches Hg.

T_f = Final absolute temperature of flask, °R.

T_i = Initial absolute temperature of flask, °R.

6.3 Sample concentration. Read μg. NO_2 for each sample from the plot of μg. NO_2 versus absorbance.

$$C=\left(\frac{m}{V_{sc}}\right)\left(\frac{1\text{ lb.}}{1.6\times10^{9}\ \frac{\mu\text{g.}}{\text{ml.}}}\right)=\left(6.2\times10^{-11}\ \frac{\text{lb./s.c.f.}}{\mu\text{g./ml.}}\right)\left(\frac{m}{V_{sc}}\right)$$

equation 7-2

where:

C = Concentration of NO_2 as NO_2 (dry basis), lb./s.c.f.

m = Mass of NO_2 in gas sample, μg.

V_{sc} = Sample volume at standard conditions (dry basis), ml.

7. *References.*

Standard Methods of Chemical Analysis. 6th ed. New York, D. Van Nostrand Co., Inc., 1962, vol. 1, p. 329–330.

Standard Method of Test for Oxides of Nitrogen in Gaseous Combustion Products (Phenoldisulfonic Acid Procedure), In: 1968 Book of ASTM Standards, Part 23, Philadelphia, Pa. 1968, ASTM Designation D–1608–60, p. 725–729.

Jacob, M. B., The Chemical Analysis of Air Pollutants, New York, N.Y., Interscience Publishers, Inc., 1960, vol. 10, p. 351–356.

Appendix B. Statistical Methods

This appendix consists of various sections which contain detailed statistical procedures carried out in the analysis of the NO_x collaborative study data. Reference to these sections has been made at various junctures in the Statistical Design and Analysis part of the body of this report. Each Appendix B section is an independent ad hoc statistical analysis pertinent to a particular problem addressed in the body of the report.

B.1 Outlier and Associated Preliminary Analysis of the Original Collaborative Test Data

After a scan of the reported collaborative test data, it was decided to base the outlier analysis on the rule: "For each collaborative test run, look at the high and low collaborator's NO_x concentration values. If such a value differs from the nearest reported value in the run by more than 10% of this nearest value, subject this suspicious outlier value to further scrutiny by recalculating its NO_x concentration from the raw reported data." For instance, on sample 14 in block 2 of the Cambridge data (cf. Table 2), the high value (1208, Lab 102) and the low value (941, Lab 103) were selected. Since 1208 differs from the nearest value (1040, Lab 104) by more than 10% $\left(\frac{1208 - 1040}{1040} = 16.2\%\right)$, it was set aside for recalculation as being a suspicious outlier point. But 941 was within 10% of its nearest value (1010, Lab 101) and was thus not subjected to calculation.

The recalculation of certain data points on the grounds that they appear to be outliers when compared against the other collaborative laboratory teams' values for the run poses a question regarding the philosophy of collaborative testing. Namely, does not such recalculation of outlier points determined by comparison with the other laboratories values for the same run tend to favorably bias an otherwise objective assessment of the precision of Method 7? The point is that, if the other collaborators' data were unavailable, one would have no inkling that these erroneous outlier points were actually calculated incorrectly. Indeed, use of the uncorrected originally reported data does actually better represent the real world situation under which performance testing for compliance using Method 7 will actually be conducted. It was decided, nevertheless, to recalculate outlier points on the rationale that with only four collaborators in the test, substantial calculation error in any outlier point could greatly, and in an ideal sense, unfairly detract from the precision of Method 7. The argument for recalculation of outlier points could be considerably enhanced through EPA's adoption of a standard Method 7 computer program by which it would process all compliance performance testing raw data, as suggested in Section III. B.

In numerous cases (those indicated by an asterisk in Table 2), calculation errors were encountered. A calculation error was defined

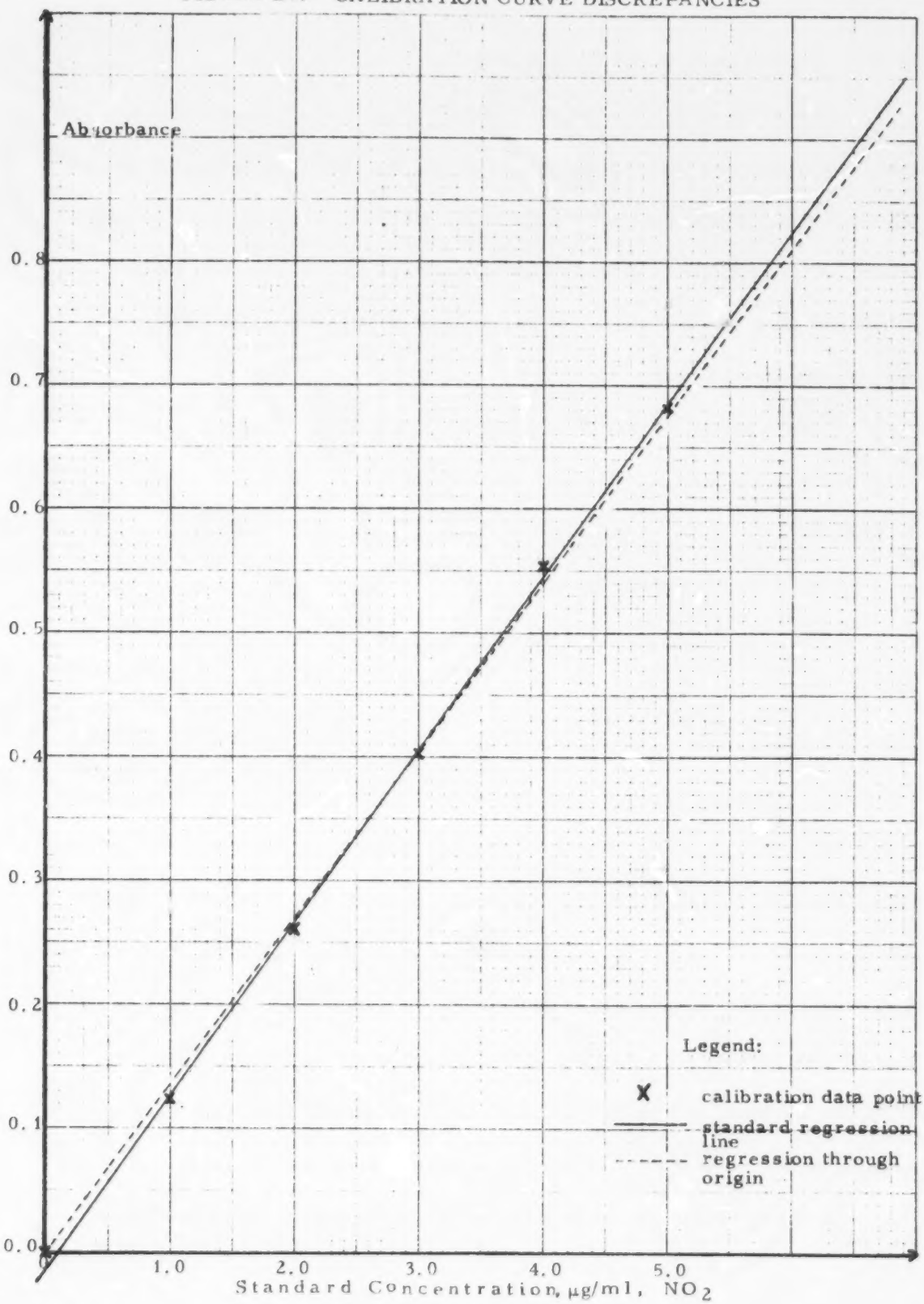
as existing when the recalculated value differed from the original reported value by more than ± 0.5% allowable for round-off differences. All of Lab 101's data was 4.84% too high because the value 6.5×10^{-5} had inadvertently been substituted for $6.2 \times 10^{-5} \frac{lb/s.c.f.}{\mu g/ml}$ in the Lab 101 computer program calculation of NO_x concentration. All of Lab 102's Cambridge data was 7.8% too high because these calculations had been performed entirely in the metric system utilizing an incorrect standard condition conversion constant to compute V_{sc}. There were also additional computational errors in the Lab 102 Cambridge NO_x concentration values for samples 12 and 20. The sample 12 error occurred somewhere in the calculation of V_{sc} while the sample 20 error was made in determining the mass of NO_2 in gas sample from the absorbance reading.

In the process of recalculating the potential outlier reported data, it was noted that a variety of procedures were utilized by the collaborative laboratories to calibrate the spectrophotometer's absorbance reading versus μg NO_2 in the sample at 420 nm. One laboratory (Lab 104) plotted the calibration data on a graph and then visually drew a curve-fitting straight line that passed through the origin. Another laboratory (Lab 103) effectively did the same by visually obtaining the slope of a straight line through the origin, differing only in that it used this slope in a computer program rather than using the calibration graph directly.

The other laboratories (Lab 101 and Lab 102) regressed their calibration data via least squares to obtain the best fitting straight line expressing absorbance as a linear function of μg/ml NO_2 in the sample. These regression lines did not pass through the origin. Figure B1 illustrates the differences in these two approaches on one of Lab 101's sets of calibration data. The approach followed by Labs 101 and 102 can introduce error into Method 7. Near the origin there can be considerable difference between the two regression lines and a large percentage error is introduced in using the standard least squares regression line. Also, the practice of diluting the sample to bring the absorbance within the range of calibration can magnify errors when an improper dilution is made that yields an absorbance either at the extreme low end or above the calibration range. It might be advisable to dictate the use of a regression line through the origin as the calibration curve and to urge that only the middle to upper portion of the calibration range be utilized whenever practicable. However, since no such statements currently exist in Method 7 as written in the Federal Register (cf. Appendix A), each laboratory's calibration curve procedures were accepted as correct in recalculating potential outlier reported data.

The corrected Cambridge data obtained as a result of the outlier analysis appear in Table 3, while the corrected Dayton data are shown in Table 4.

FIGURE B1. CALIBRATION CURVE DISCREPANCIES



B.2 Significance of the Port Effect

Because of the limited number of ports available on a stack, each collaborative team is constrained to a sample from only one port on each run. Since there is likely to be non-uniform flow through the stack, a potential port effect is introduced. This section determines whether there was a significant port effect in either the Cambridge test or the Dayton test.

The statistical test utilized to assess the significance of the port main effect was the unsophisticated rank test that Youden presented in his address, "The Collaborative Test."[1] The null hypothesis is that there is no difference between the n = 4 ports. Each of the M = 16 runs at a test site are considered different materials in the Youden terminology. Ranks from 1 to 4 are assigned to the ports' data for each of the 16 runs. These sixteen ranks for a port are summed to yield its score. The approximate .05 level lower and upper limits for the two sided test are scores of 28 and 52, respectively. On the Cambridge test data, the low score was 35.5 for port A; ports B and D had the highest scores of 42.5. On the Dayton test data, port D had the lowest score (36.5) while port B had the highest score (41.5). Clearly, this test detects no significant port effect in either the Cambridge or the Dayton test data. It should be noted the existence of a sizable collaborator main effect would, by the very nature of the randomized block design employed, tend to obscure any small port effect that might be present. Thus, while a

port effect may in fact exist, its influence on the data is undetectable and hence, in effect, negligible. Therefore, one can proceed with the precision statistical analysis on the tenable assumption that there is no port effect.

B.3 Empirical Relationship of the Standard Deviation to the Mean in the Collaborative Test Data

A question that is fundamental to the proper statistical analysis of the collaborative test data is whether, and to what extent, the variation in the Method 7 NO_x concentration values, both within a single laboratory and between the collaborative laboratories, is related to the true concentration level of the measurement. To phrase the question in statistical terminology, is there a functional relationship of the inter-laboratory (i.e., between laboratory) run standard deviation to the run mean and of the intra-laboratory (i.e., within laboratory) collaborator block standard deviation to the collaborator block mean?

Let x_{ij} denote the Method 7 NO_x concentration level reported by laboratory i ($i = 1, 2, 3, 4$) on run j ($j = 1, 2, \ldots, 16$) with runs 1 through 4 comprising block 1, runs 5 through 8 in block 2, etc. Then the inter(between)-laboratory run j standard deviation estimate may be defined as $S_j = \sqrt{\frac{1}{3} \sum_{i=1}^{4} (x_{ij} - \bar{x}_{\cdot j})^2}$ where the inter-laboratory run j mean estimate $\bar{x}_{\cdot j} = \frac{1}{4} \sum_{i=1}^{4} x_{ij}$. The inter-laboratory block k mean estimate is defined as $\bar{x}_{\cdot k} = \frac{1}{4} \sum_{j=4k-3}^{4k} \bar{x}_{\cdot j}$. Table B1 contains

TABLE B1. INTER-LABORATORY RUN SUMMARY

Test Site	Block k	Block Mean $\bar{\bar{x}}_{\cdot k}$ 10^{-7} lb/scf	Run j	Mean $\bar{\bar{x}}_{\cdot j}$ 10^{-7} lb/s.c.f.	Std. Dev. S_j lb/s.c.f.	Coefficient of Variation $b_j = S_j/\bar{\bar{x}}_{\cdot j}$
Cambridge						
	1	1455	1	1394	59.0	.0423
			2	1520	200.2	.1317
			3	1464	44.0	.0301
			4	1442	67.3	.0467
	2	1004	5	1110	144.7	.1303
			6	962	115.8	.1204
			7	1016	81.3	.0800
			8	926	179.0	.1932
	3	679	9	651	148.7	.2283
			10	651	90.7	.1394
			11	704	68.3	.0973
			12	709	20.5	.0289
	4	385	13	383	33.9	.0880
			14	394	37.7	.0957
			15	374	11.3	.0302
			16	389	26.4	.0679
Dayton						
	1	466	1	454	34.9	.0769
			2	494	27.7	.0561
			3	479	14.2	.0296
			4	439	38.2	.0870
	2	356	5	338	21.1	.0624
			6	338	35.1	.1038
			7	379	33.5	.0885
			8	369	30.6	.0828
	3	224	9	229	16.0	.0699
			10	234	14.9	.0638
			11	222	7.4	.0333
			12	213	18.0	.0845
	4	122	13	127	11.7	.0925
			14	120	12.6	.1053
			15	124	11.0	.0889
			16	119	14.3	.1202

the inter-laboratory block mean and run mean and standard deviation estimates for both the Cambridge and Dayton collaborative test data. The inter-laboratory run standard deviation is plotted as a function of the inter-laboratory run mean in Figure B2. The extent of scatter in this data reflects the limitation of having only four laboratories sampling simultaneously. From this figure it is quite evident that the run standard deviation increases as the run mean increases. Thus a stepwise multiple linear regression program was utilized to determine the best least squares regression line passing through the origin to fit the Figure B2 data. Regressor variables provided to fit the standard deviation S_j as a function of the mean $\overline{x}$ included $\overline{x}^{-1}$, $\overline{x}^{1/4}$, $\overline{x}^{1/2}$, $\overline{x}^{3/4}$, $\overline{x}$, $\overline{x}^{5/4}$, $\overline{x}^{3/2}$, $\overline{x}^{2}$, and $\log \overline{x}$. It was found that the linear function $S_j = b\overline{x}_{\cdot j}$ provided a better fit than any other simple or multiple regression. Even though there is much scatter in the data, the linear regression $S_j = b\overline{x}_{\cdot j}$ strongly suggests that the inter-laboratory run standard deviation is directly proportional to the average inter-laboratory run NO_x concentration.

A similar analysis can be performed on the intra(within)-laboratory block statistics to see what relationship holds at the intra-laboratory level. Table B2 presents the intra-laboratory collaborator block means $\overline{x}_{i\cdot k} = \frac{1}{4}\sum_{j=4k-3}^{4k} x_{ij}$ and standard deviations $S_{ik} = \sqrt{\frac{1}{3}\sum_{j=4k-3}^{4k}(x_{ij} - \overline{x}_{i\cdot k})^2}$ for each collaborator i in each block k at both the Cambridge and Dayton

FIGURE B2. INTER-LABORATORY RUN PLOT

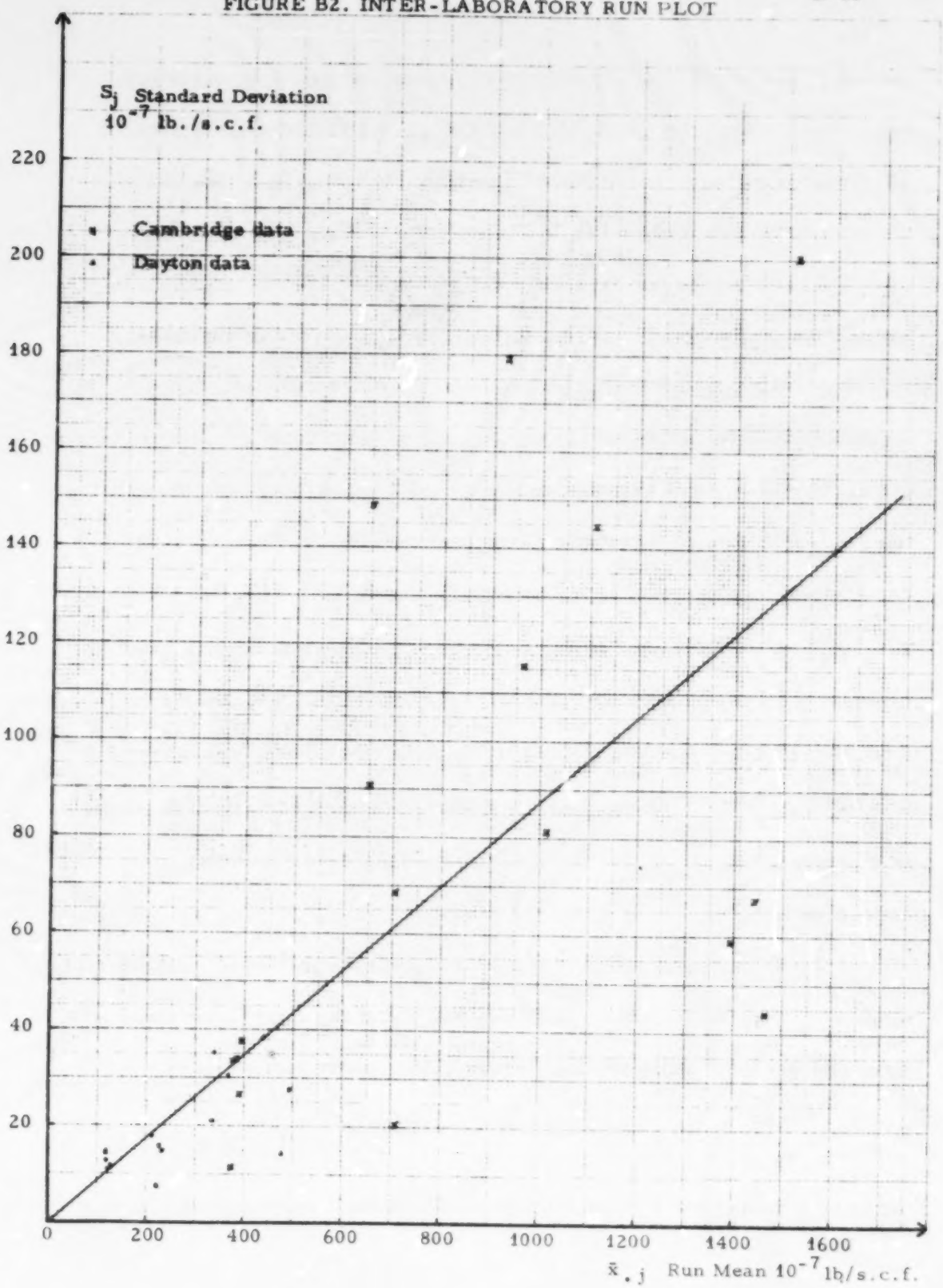


TABLE B2. INTRA-LABORATORY COLLABORATOR BLOCK SUMMARY

Test Site	Intra-Laboratory Collaborator Block				
	Block k	Collaborator i	Mean $\bar{x}_{i \cdot k}$ 10^{-7} lb/s.c.f.	Std. Dev. S_{ik}	Coef. of Variation $b_{ik} = S_{ik}/\bar{x}_{i \cdot k}$
Cambridge					
	1	Lab 101	1459	27.6	.0190
		Lab 102	1447	81.2	.0561
		Lab 103	1532	192.6	.1257
		Lab 104	1383	37.8	.0273
	2	Lab 101	855	154.5	.1807
		Lab 102	1138	128.8	.1132
		Lab 103	994	40.7	.0409
		Lab 104	1027	49.9	.0486
	3	Lab 101	583	122.0	.2095
		Lab 102	703	40.2	.0571
		Lab 103	754	44.9	.0596
		Lab 104	675	17.3	.0257
	4	Lab 101	361	12.3	.0342
		Lab 102	392	15.1	.0386
		Lab 103	417	25.7	.0617
		Lab 104	370	11.6	.0312
Dayton					
	1	Lab 101	453	14.2	.0314
		Lab 102	444	44.6	.1004
		Lab 103	503	17.4	.0345
		Lab 104	465	28.9	.0621
	2	Lab 101	333	13.8	.0415
		Lab 102	330	27.3	.0827
		Lab 103	391	28.0	.0716
		Lab 104	370	18.3	.0493
	3	Lab 101	217	5.3	.0245
		Lab 102	233	12.4	.0534
		Lab 103	211	15.9	.0754
		Lab 104	237	9.6	.0403
	4	Lab 101	117	2.9	.0252
		Lab 102	134	5.6	.0418
		Lab 103	108	7.6	.0705
		Lab 104	130	0.0	.0000

test sites. The intra-laboratory collaborator block standard deviation is plotted as a function of the intra-laboratory collaborator block mean in Figure B3. Again a definite linear relationship is evident. Stepwise regression through the origin with the same regressor variables as above showed that the linear relationship $S_{ik} = b\bar{x}_{i\cdot k}$ gave a better fit to the data than did any other simple or multiple regression. The selection of the linear regression $S_{ik} = b\bar{x}_{i\cdot k}$ indicates that the intra-laboratory collaborator block standard deviation is also proportional to the collaborator block mean.

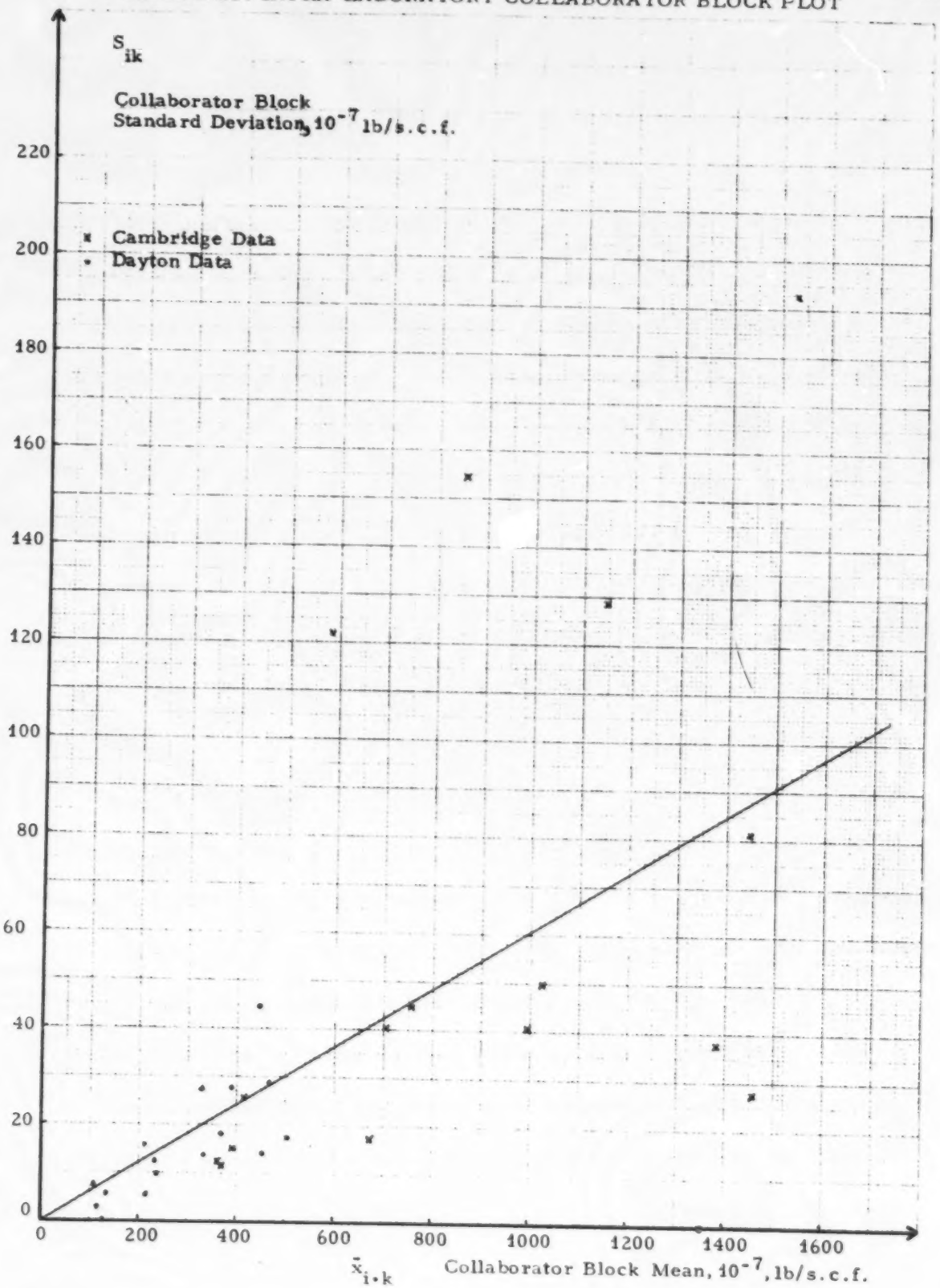
B.4 Transformation of the Collaborative Test Data

To answer the question of what is the appropriate data transformation to achieve run equality of variance over both the Cambridge and the Dayton collaborative test data, several likely such transformations were selected and evaluated using Bartlett's test.[2] These transformations were the linear (no transformation of x), the logarithmic ($\log_{10}x$), and the square root ($x^{1/2}$). Table B3 contains the results of this transformation run equality of variance analysis. Since Bartlett's test has the

TABLE B3. DATA TRANSFORMATIONS TO ACHIEVE RUN EQUALITY OF VARIANCE

Transformation		Bartlett's Test Statistic	Significance
Linear	$y = x$	$B = 128.7$	$P\{\chi^2(31) \geq 128.7\} << .0001$
Logarithmic	$y = \log_{10}(x)$	$B = 47.7$	$P\{\chi^2(31) \geq 47.7\} = .029$
Square root	$y = x^{1/2}$	$B = 74.9$	$P\{\chi^2(31) \geq 74.9\} < .0001$

FIGURE B3. INTRA-LABORATORY COLLABORATOR BLOCK PLOT



null hypothesis of variance equality, it is only for a small value of the test statistic B that this hypothesis should not be rejected. Clearly the linear and square root transformations are inferior to the logarithmic transformation. However, even under the logarithmic transformation, there is only a 2.9% chance that true equality of run variance would have produced data with as much scatter in the run variances as that actually obtained in the collaborative test data. Still, the logarithmic transformation is the best transformation available and it is far superior to no data transformation at all.

It was on the above rationale that the logarithmic transformation was used in substituting a value for the missing Lab 101 measurement through port C in run 4 (sample 11) in the first block of the Cambridge collaborative test data. All the data in this block were transformed using the $\log_{10}$ transformation. Then the minimum variance unbiased estimate of the missing Lab 101 measurement was obtained by adding Lab 101's net difference from the other three runs' means to the run 4 mean. This gave a value of 3.1603 for this missing point in the logarithmic scale which translates to 1446×10^{-7} lb/s.c.f. in the original linear scale.

It can be shown that acceptance of the data transformation log x based upon (run) variance equality considerations is equivalent to acceptance of the direct proportionality of the (run) standard deviation to the (run) mean. The proof is as follows. Suppose X follows a lognormal

distribution with mean μ_x and variance σ_x^2. Then log X has a normal distribution with mean $\mu_{\log x}$ and variance $\sigma_{\log x}^2$. Now the mean and variance of X can be expressed in terms of the logarithmic transformed mean $\mu_{\log x}$ and variance $\sigma^2_{\log x}$: (3)

$$\mu_x = e^{(\mu_{\log x} + \frac{1}{2}\sigma^2_{\log x})}$$

$$\sigma^2_x = (e^{\sigma^2_{\log x}} - 1)\left[e^{(\mu_{\log x} + \frac{1}{2}\sigma^2_{\log x})}\right]^2$$

Substituting, $$\sigma^2_x = (e^{\sigma^2_{\log x}} - 1)\left[\mu_x\right]^2$$

$$e^{\sigma^2_{\log x}} = 1 + \left(\frac{\sigma_x}{\mu_x}\right)^2$$

Thus, $$\sigma^2_{\log x} = \log\left[1 + \left(\frac{\sigma_x}{\mu_x}\right)^2\right]$$

Now when there is equality of variance under the logarithmic transformation, $\sigma^2_{\log x}$ = constant, say C_1.

$$\log\left[1 + \left(\frac{\sigma_x}{\mu_x}\right)^2\right] = \sigma^2_{\log x} = C_1$$

$$1 + \left(\frac{\sigma_x}{\mu_x}\right)^2 = C_2$$

So $$\frac{\sigma_x}{\mu_x} = \text{constant, say } \beta.$$

Therefore $$\sigma_x = \beta\mu_x$$

This demonstrates that when there is equality of variance under the logarithmic transformation, then the standard deviation is proportional to the mean. Therefore, the selection of the logarithmic transformation above based upon equality of the inter-laboratory run variances under the logarithmic transformation is equivalent to a demonstration of the proportionality of the inter-laboratory run standard deviation to the run mean.

B.5 The Underlying Relationship of the Standard Deviation Components to the Mean

From Appendix B.3 it was shown that the Cambridge and Dayton collaborative test data exhibit a linear relationship of the intra-laboratory collaborator block sample standard deviation to the sample mean:

$$S_{ik} = b\bar{x}_{i \cdot k}$$

where b is the sample coefficient of variation. Now, the four sample measurements made by a collaborator within each block are the closest possible approximation to replicates, given the randomly varying nature of NO_x emissions in the stack of a stationary emissions source. In fact, these four collaborator block measurements duplicate the circumstances encountered in making the three field measurements that constitute the performance test for compliance. Hence, the expected value of the collaborator block sample standard deviation is σ, the within laboratory standard deviation of replicate measurements. The sample standard deviation is a biased estimate of the population standard deviation in an

underlying normal distribution.[4] For a sample size of n = 4 replicates, the correction factor is $\alpha_4 = \sqrt{\frac{2}{3}} \frac{\Gamma(4/2)}{\Gamma[(4-1)/2]} = .92132 = \frac{1}{1.0854}$:

$$E(1.0854\, S_{ik}) = \sigma$$

So $\sigma = 1.0854 E(S_{ik}) = 1.0854\, E(b\bar{x}_{i.k}) = 1.0854\, b\, E(\bar{x}_{i.k}) = 1.0854\, b\mu$

Define the within lab coefficient of variation as $\beta = 1.0854b$. Then the within lab standard deviation is given by

$$\sigma = \beta\mu$$

The inter-laboratory run sample standard deviation was shown in both Appendix B.3 and Appendix B.4 to be proportional to the sample mean, with a sample coefficient of variation, say c:

$$S_j = c\bar{x}_{.j}$$

Now the variance in the four collaborators' measurements for a run has both a within laboratory variance component σ^2 and a true lab mean to true lab mean variance component σ_L^2. Thus

$$E(S_j^2) = \sigma^2 + \sigma_L^2$$

Accounting for a sample size of n = 4 collaborators, the expected value of $1.0854\, S_j$ is

$$E(1.0854\, S_j) = \sqrt{E(S_j^2)} = \sqrt{\sigma^2 + \sigma_L^2}$$

So $\sqrt{\sigma^2 + \sigma_L^2} = 1.0854\, E(S_j) = 1.0854\, c\, E(\bar{x}_{.j}) = 1.0854\, c\mu$

$$\sigma_L^2 = (1.0854\, c\mu)^2 - \sigma^2$$

Substituting from the intra-laboratory relationship above,

$$\sigma_L^2 = (1.0854\, c\mu)^2 - (1.0854\, b\mu)^2$$

$$\sigma_L = \left(1.0854\ \sqrt{c^2 - b^2}\right)\mu$$

Define the laboratory coefficient of variation $\beta_L = \left(1.0854\ \sqrt{c^2 - b^2}\right)$.

Then

$$\sigma_L = \beta_L \mu$$

B.6 Estimating the Standard Deviation Components and the Reproducibility Standard Deviation

The preceding sections have demonstrated that the within laboratory standard deviation component σ and the true lab mean-to-true lab mean standard deviation component σ_L both obey the coefficient of variation hypothesis: $\sigma = \beta\mu$ and $\sigma_L = \beta_L \mu$. In the coefficient of variation situation, Zeigler[5] has recently shown that the best estimator of β for values of $\beta < .2$ is

$$\hat{\beta} = \frac{C_n}{K} \sum_{i=1}^{K} b_i$$

where $b_i = \frac{S_i}{\bar{x}_i}$, $E(C_n S) = \sigma$, and K is the number of samples of size n. For our design, n = 4 and K = 32; so $C_n = 1.0854$.

Thus $$\hat{\beta} = \frac{1.0854}{32} \sum_{i=1}^{32} \frac{S_i}{\bar{x}_i}.$$

This estimator provides the basis for estimating σ and σ_L for Method 7.

Let us first consider the within lab standard deviation σ.

Let

$$\hat{\beta} = \frac{1.0854}{32} \sum_{ik=1}^{32} \frac{S_{ik}}{\bar{x}_{i\cdot k}}$$

where the index ik is taken over the 32 collaborator block groups of data at Cambridge and Dayton. The actual individual values $S_{ik}/\bar{x}_{i\cdot k}$ are presented as the last column in Table B2. Their average multiplied by 1.0854 yields

$$\hat{\beta} = 1.0854(.06042) = .06558$$

Deriving the expected value of $\hat{\beta}$:

$$E(\hat{\beta}) = E\left[\frac{1.0854}{K} \sum_{ik=1}^{K} \frac{S_{ik}}{\bar{x}_{i\cdot k}}\right]$$

$$E(\hat{\beta}) = \frac{1}{K} \sum_{ik=1}^{K} E\left(\frac{1.0854\, S_{ik}}{\bar{x}_{i\cdot k}}\right)$$

Now $E\left(\frac{S}{\bar{x}}\right) \approx \frac{E(S)}{E(\bar{x})}$. Within each block the measurements are approximately normally destributed. Under the normal distribution, the mean and variance are independent. Hence, the sample mean and sample standard deviation should be nearly independent. Thus the expected value of their ratio is approximated by the ratio of their expected values.

So

$$E(\hat{\beta}) \approx \frac{1}{K} \sum_{ik=1}^{K} \frac{E(1.0854\, S_{ik})}{E(\bar{x})}$$

$$= \frac{1}{K} \sum_{ik=1}^{K} \frac{\sigma}{\mu} \quad \text{from Appendix B.5.}$$

$$E(\hat{\beta}) \approx \frac{\sigma}{\mu}$$

since $\hat{\beta} = .06558$, $\hat{\sigma} = \hat{\beta}\mu = .06558\mu$. Thus for Method 7, the within laboratory standard deviation is 6.558% of the mean value. According to Mandel, the degrees of firmness in this within laboratory estimate over g = 8 blocks is

$$\phi = pg(n - 1) = 4 \cdot 8(4 - 1) = 96$$

The percentage coefficient of variation of the within lab standard deviation estimate is thus

$$\%CV_{\hat{\sigma}} = 100\sqrt{\frac{1}{2\phi}} = 7.2\%$$

Next consider the standard deviation component σ_L due to true laboratory-to-laboratory variation. In this case, the estimated value of the between-laboratory coefficient of variation β_b is

$$\hat{\beta}_b = \frac{1.0854}{32} \sum_{j=1}^{32} \frac{S_j}{\bar{x}_{\cdot j}}$$

where the run index j extends over the 32 collaborative test runs at Cambridge and Dayton. The actual $S_j / \bar{x}_{\cdot j}$ values are tabulated in the last column of Table B1. The estimated value of β_b is

$$\hat{\beta}_b = 1.0854(.08739) = .09485$$

The expected value of β_b is derived below.

$$E(\hat{\beta}_b) = \frac{1}{K} \sum_{j=1}^{K} E\left(\frac{1.0854 S_j}{\bar{x}_{\cdot j}}\right)$$

$$\approx \frac{1}{K} \sum_{j=1}^{K} \frac{E(1.0854 S_j)}{E(\bar{x}_{\cdot j})}$$ under the rationale presented above.

$$= \frac{1}{K} \sum_{j=1}^{K} \frac{\sqrt{\sigma^2 + \sigma_L^2}}{\mu} \quad \text{from Appendix B.5.}$$

Thus

$$E(\hat{\beta}_b) \approx \frac{\sqrt{\sigma^2 + \sigma_L^2}}{\mu}$$

$$\sqrt{\hat{\sigma}^2 + \hat{\sigma}_L^2} = \hat{\beta}_b \mu$$

$$\hat{\sigma}_L = \mu \sqrt{\hat{\beta}_b^2 - \hat{\beta}^2} \quad \text{substituting } \hat{\beta}\mu \text{ for } \hat{\sigma}.$$

$$\hat{\sigma}_L = \mu \sqrt{(.09485)^2 - (.06558)^2}$$

$$\hat{\sigma}_L = .06853\mu$$

For Method 7, the true laboratory-to-laboratory standard deviation component is 6.853% of the mean value.

Mandel defines reproducibility in terms of the results of a performance test for compliance. The reproducibility standard deviation estimate is defined as $\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/m}$ where m is the number of replicate measurements comprising the test result. For the Method 7 test result, m is usually 12. Thus the reproducibility standard deviation for Method 7 is given by

$$\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/12} = \sqrt{(.06853\mu)^2 + (.06558\mu)^2/12} = .07110\mu$$

Therefore the Method 7 test result reproducibility standard deviation is 7.11% of the result. Mandel's reproducibility degrees of firmness is given below, where $\hat{\gamma} = \frac{\hat{\sigma}_L^2}{\hat{\sigma}^2} = \frac{(.06853\mu)^2}{(.06558\mu)^2} = 1.092$:

$$\phi = \frac{(ng)^2\left(\gamma + \frac{1}{m}\right)^2}{\frac{(ng - m)^2}{pg(n-1)m^2} + \frac{(1 + ng\gamma)^2}{p - 1}} = 3.284$$ for the randomized block

design with p = 4 laboratories making n = 4 replicate measurements in g = 8 blocks. The percentage coefficient of variation of the reproducibility standard deviation estimate is

$$\frac{\%CV}{\sqrt{\hat{\sigma}_L^2 + \hat{\sigma}^2/12}} = 100\sqrt{\frac{1}{2\phi}} = 39.0\%$$

B.7 The Nitrate Solution Data

The nitrate solution data of all four collaborative laboratories analyzed in conjunction with both the Cambridge and the Dayton test samples is presented in Table B.4. As instructed, the data are reported in units of μg NO_2 per 10 ml of absorbance sample. The instructions for analyzing these nitrate solutions and the data reporting form are shown in Figure 8. The Table B4 data averaged over replicates and days are shown in Table B5. The Table B6 data are further averaged over site samples. It is interesting to note in Table B5 that for Solutions A, B, and C, the average values for samples analyzed along with Cambridge test samples are consistently higher than the corresponding average values obtained in the Dayton sample analysis. Another noteworthy characteristic is the difference in laboratory means shown in Table B6 for the collaborative laboratories. For each of the three actual nitrate

TABLE B4. REPORTED NITRATE SOLUTION CONCENTRATIONS, μg NO_2 per 10 ml

Collaborative Laboratory	Day	Repl.	With Cambridge Samples Sol.A	Sol.B	Sol.C	Sol.D	With Dayton Samples Sol.A	Sol.B	Sol.C	Sol.D
Lab 101	1	1	22.31	11.13	35.58	1.71	23.21	11.80	36.19	0.68
		2	21.18	11.13	36.50	1.08	22.85	11.44	36.05	0.39
		3	22.21	11.79	36.50	2.11	23.28	11.08	35.98	0.68
	2	1	25.18	11.79	38.39	2.11	23.16	11.29	34.60	0.48
		2	25.50	13.10	37.89	2.11	23.16	11.01	34.95	0.36
		3	25.18	13.55	37.89	2.11	23.45	11.01	34.60	0.43
	3	1	24.84	12.99	37.03	2.15	24.20	11.78	36.40	-0.21
		2	25.36	12.99	37.66	1.61	24.13	12.06	36.54	-0.14
		3	25.36	12.99	38.28	2.15	24.05	12.13	36.33	0.43
Lab 102	1	1	26.64	12.97	34.06	0.	23.44	12.66	37.66	0.657
		2	22.03	13.20	39.53	0.	25.66	14.23	34.30	1.086
		3	26.64	10.62	34.06	0.	24.87	13.66	34.09	1.014
	2	1	29.40	16.15	37.09	2.48	24.23	13.59	34.16	1.657
		2	28.55	15.73	37.52	2.74	26.30	13.66	37.23	1.871
		3	29.40	15.98	37.09	3.33	23.66	12.73	39.66	1.657
	3	1	27.46	11.78	39.32	0.	20.80	12.37	40.66	2.510
		2	27.46	11.78	40.17	0.	22.59	13.30	38.30	1.657
		3	26.61	12.20	41.86	0.	21.73	13.66	34.80	1.942
Lab 103	1	1	27.9	13.3	40.0	1.0	26.6	13.3	40.9	0.83
		2	27.1	13.8	39.2	0.4	28.2	14.2	40.9	0.83
		3	27.1	14.2	39.2	0.4	25.0	13.3	38.4	0.83
	2	1	27.2	13.6	40.0	0.	23.3	14.2	38.4	0.
		2	26.8	12.3	39.2	0.	27.5	12.5	41.6	0.
		3	27.1	13.9	40.0	0.	27.5	13.3	38.4	0.
	3	1	27.1	13.5	41.3	0.	27.5	12.5	40.0	0.83
		2	26.1	14.2	41.3	0.	26.6	13.3	41.6	0.
		3	26.8	12.7	39.2	0.	27.5	14.2	39.2	0.
Lab 104	1	1	26.6	13.4	41.0	0.9	24.8	13.7	39.9	1.0
		2	27.0	13.4	40.2	0.8	26.7	12.7	37.0	0.6
		3	27.0	13.5	40.7	0.7	26.6	14.1	40.0	0.5
	2	1	26.0	14.6	39.9	0.6	26.0	12.7	38.6	0.4
		2	26.2	13.4	39.9	0.8	24.4	12.9	37.0	0.4
		3	25.9	13.6	39.9	0.8	25.8	13.0	38.3	0.4
	3	1	25.4	13.4	38.3	0.5	25.1	13.0	34.0	0.6
		2	25.7	12.7	38.8	0.6	24.4	13.3	34.7	0.6
		3	24.8	12.6	36.2	0.6	25.8	12.8	36.2	0.8

TABLE B5. DAY AVERAGED NITRATE SOLUTION CONCENTRATIONS, $\mu g\ NO_2$ per 10 ml

Collaborative Laboratory	With Cambridge Samples				With Dayton Samples			
	Sol. A	Sol. B	Sol. C	Sol. D	Sol. A	Sol. B	Sol. C	Sol. D
Lab 101	24.12	12.38	37.30	1.90	23.50	11.51	35.74	0.34
Lab 102	27.13	13.37	37.86	0.95	23.70	13.31	36.76	1.56
Lab 103	27.02	13.50	39.93	0.20	26.63	13.42	39.93	0.37
Lab 104	26.07	13.40	39.43	0.70	25.51	13.13	37.30	0.59

TABLE B6. AVERAGE LABORATORY NITRATE SOLUTION CONCENTRATION, μg NO_2 per 10 ml

Collaborative Laboratory	Solution A	Solution B	Solution C	Solution D
Lab 101	23.81	11.95	36.52	1.12
Lab 102	25.42	13.35	37.31	1.25
Lab 103	26.83	13.46	39.93	0.28
Lab 104	25.79	13.27	38.37	0.64

solutions (A, B, and C), the Lab 103 average was highest, while the Lab 101 average was lowest.

B.8 The Variance Components of the Analytical Procedure

A separate analysis of variance was performed on the data for each of the four nitrate solutions presented in Table B4. The crossed factors were collaborator C and site S from which the concurrently analyzed test samples came. The day factor D(CS) was nested within C x S, and the replicate factor R(CSD) was nested within days. The resulting analyses and components of variance for a random effects model are presented in Table B7. The F-ratios and significance tests for the factors are given in Table B8. With all four solutions, Table B8 shows that the day factor D(CS) is very significant ($P \leq .002$), while the collaborator factor C, the test sample source site S, and the collaborator-site interaction factor CS are never significant ($P > .05$). Therefore, the only important variance components are those for the day factor σ^2_D and the replication (within lab) factor σ^2_R. Estimates of these variance components are given in Table B7 for each of the four solutions.

For comparison with collaborative test data variance component estimates, $\hat{\sigma}^2$ and $\hat{\sigma}^2_L$ for Method 7, it is necessary to estimate the the corresponding within lab and laboratory-to-laboratory variance components for the analytical portion of Method 7 from the nitrate

TABLE B7. NITRATE SOLUTION DATA ANALYSES OF VARIANCE

Factor	Sum of Squares	D.F.	Mean Square	Expected Mean Square	Variance Component
Solution D	Mean = 0.827				
C	10.798	3	3.599	$18\sigma_C^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_C^2 = 0$
S	.893	1	.893	$36\sigma_S^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_S^2 = 0$
CS	11.923	3	3.974	$9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_{CS}^2 = .2948$
D(CS)	21.137	16	1.321	$3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_D^2 = .4205$
R(CSD)	2.855	48	.060	σ_R^2	$\hat{\sigma}_R^2 = .0595$
Solution B	Mean = 13.006				
C	27.218	3	9.073	$18\sigma_C^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_C^2 = .3690$
S	1.840	1	1.840	$36\sigma_S^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_S^2 = 0$
CS	1.956	3	.652	$9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_{CS}^2 = 0$
D(CS)	38.890	16	2.431	$3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_D^2 = .6841$
R(CSD)	18.153	48	.378	σ_R^2	$\hat{\sigma}_R^2 = .3782$
Solution A	Mean = 25.459				
C	84.894	3	28.298	$18\sigma_C^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_C^2 = 1.0388$
S	28.050	1	28.050	$36\sigma_S^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_S^2 = .5125$
CS	28.798	3	9.599	$9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_{CS}^2 = .5562$
D(CS)	73.498	16	4.594	$3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_D^2 = 1.1970$
R(CSD)	48.124	48	1.003	σ_R^2	$\hat{\sigma}_R^2 = 1.0026$
Solution C	Mean 38.032				
C	117.650	3	39.217	$18\sigma_C^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_C^2 = 1.3580$
S	25.824	1	25.824	$36\sigma_S^2 + 9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_S^2 = .5501$
CS	11.049	3	3.683	$9\sigma_{CS}^2 + 3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_{CS}^2 = 0$
D(CS)	96.325	16	6.020	$3\sigma_D^2 + \sigma_R^2$	$\hat{\sigma}_D^2 = 1.331$
R(CSD)	97.346	48	2.028	σ_R^2	$\hat{\sigma}_R^2 = 2.028$

TABLE B8. SIGNIFICANCE OF NITRATE SOLUTION FACTORS

Solution	Factor	F-ratio	Significance
D			
	C	0.906	P > .50
	S	0.225	P > .50
	CS	3.008	P = .07
	D(CS)	22.202	P < .001
B			
	C	3.732	P ≈ .15
	S	0.757	P > .50
	CS	0.268	P > .50
	D(CS)	6.431	P < .001
A			
	C	2.948	P ≈ .18
	S	2.922	P ≈ .19
	CS	2.089	P ≈ .15
	D(CS)	4.580	P < .001
C			
	C	6.514	P = .08
	S	4.290	P ≈ .13
	CS	0.612	P > .50
	D(CS)	2.968	P ≈ .002

solution data. Now, in both the Method 7 performance test for compliance and in the Method 7 collaborative test situations, the term repeatability refers to the replicate measurement variation in samples collected at specific 20 to 30 minute time intervals on a single day and subsequently analyzed together in the laboratory on a later single day. The single day for sample laboratory analysis becomes an important consideration as a result of the preceding analyses of variance showing that the fluctuation from day-to-day in each laboratories' analyses was very significantly larger than the single day replicate analysis fluctuation. Consequently, the appropriate comparable definition of the within lab variance component σ^2 in the nitrate solution analysis situation is $\sigma^2 = \sigma^2_R$, the replication variance component in a single day's repeated measurements.

Again, in both the Method 7 performance test for compliance and in the Method 7 collaborative test situation, the laboratory-to-laboratory variance component σ^2_L of the reproducibility refers to samples collected simultaneously, either hypothetically or actually, by various laboratory teams at the test site for subsequent analysis at a later date in their respective laboratories. Note the day factor is here effectively nested within the laboratory factor by the fact that the different laboratories perform their sample analyses on different days, in general, under different environmental conditions, each requiring a different calibration of the spectrophotometer. Thus

rather than naively assume that collaborator variance component σ^2_C from the preceding analyses of variance is synonymous with σ^2_L, one must instead take into account the major effect of the day variance σ^2_D and the occasional lesser effects of σ^2_S and σ^2_{CS} in estimating σ^2_L. The best approach is to duplicate the design situation and the calculation method by which σ^2_L was estimated from the collaborative test data for estimating σ^2_L for the nitrate solution data. Consequently, the between-laboratory mean squares were computed as

$$MS_L = \frac{1}{18} \sum_{S,d,r=1}^{18} \left[\frac{1}{3} \sum_{C=1}^{4} (x_{CSdr} - \bar{x}_{.Sdr})^2 \right]$$

for each of the four solutions, at each of the 18 site, day, replicate combinations. Since the expected value of MS_L is $E(MS_L) = \sigma^2_L + \sigma^2$, one can estimate $\hat{\sigma}^2_L$ for each solution, using the estimate of $\hat{\sigma}^2 = \hat{\sigma}^2_R$ from its analysis of variance given in Table B7. The values of MS_L and $\hat{\sigma}^2_L$ so obtained are presented in Table B9 along with a summary of all the pertinent variance and standard deviation components for the four nitrate solutions. Note that the nitrate solution data indicate that the analytical laboratory standard deviation components $\hat{\sigma}$ and $\hat{\sigma}_L$ are <u>not</u> proportional to the mean μ, as the collaborative test data have shown Method 7 to be. Still the analytical $\hat{\sigma}_L$, and particularly $\hat{\sigma}$, do appear to be linear functions of μ. Linear

TABLE B9. NITRATE SOLUTION VARIANCE AND STANDARD DEVIATION COMPONENTS

	Sol. D	Sol. B	Sol. A	Sol. C
μ, $\mu g\ NO_2/10ml$	0.00	12.50	25.00	37.50
MS_L	.7689	1.3022	3.8035	5.4684
Repeatability Variation				
$\hat{\sigma}^2 = \hat{\sigma}^2_R$	.0595	.3782	1.0026	2.0280
$\hat{\sigma}$, $\mu g\ NO_2/10\ ml$	.2439	.6150	1.0013	1.4241
Lab-to-lab Variation				
$\hat{\sigma}^2_L = MS_L - \hat{\sigma}^2$	.7094	.9240	2.8009	3.4404
$\hat{\sigma}_L$, $\mu g\ NO_2/10\ ml$	.8423	.9612	1.6736	1.8548

regression yields the following mean level dependent standard deviation component equations for the analytical laboratory phase:

$$\hat{\sigma} = .03142\mu + .2321\mu g\ NO_2/10\ ml$$

$$\hat{\sigma}_L = .03000\mu + .7705\mu g\ NO_2/10\ ml$$

B.9 The Minimum Detectable Limit

The minimum detectable limit of Method 7 is the smallest Method 7 performance test result that is significantly larger than zero in the reproducibility sense. In Table 9 of the Sources of Variability section, it is shown that nearly all of the variability in Method 7 at very low absorbance sample NO_2 concentrations is due to the analytical portion of Method 7. Hence, at zero NO_2 concentration, we will assume that all of the variability results from the analytical portion. Now, Solution D of the nitrate solution did, in fact, represent a zero nitrogen oxide emission level. From Table B9, the within laboratory standard deviation calculated for Solution D is $\hat{\sigma} = .2439\mu g\ NO_2/10\ ml$ and the lab-to-lab standard deviation for Solution D is $\hat{\sigma}_L = .8423\mu g\ NO_2/10\ ml$. Since an absorbance sample of 100 ml is specified in Method 7, the standard deviation estimates at zero NO_x concentration expressed in mass units are:

$$\hat{\sigma} = (.2439\mu g\ NO_2/10\ ml)(100\ ml) = 2.439\mu g\ NO_2$$

$$\hat{\sigma}_L = (.8423\mu g\ NO_2/10\ ml)(100\ ml) = 8.423\mu g\ NO_2$$

The reproducibility standard deviation at zero NO_2 concentration is estimated as $\sqrt{\hat{\sigma}^2_L + \hat{\sigma}^2/12} = 8.452 \mu g\ NO_2$. Assuming a sample volume of 2000 ml is collected at standard conditions, the upper 95% confidence limit for a true concentration $\mu_C = 0$ is

$$C_{.95} = \left(620 \times 10^{-7} \frac{lb/s.c.f.}{\mu g/ml}\right) \frac{1.96\sqrt{\hat{\sigma}_L + \hat{\sigma}^2/12}}{V_{sc}}$$

$$= \frac{(620 \times 10^{-7})(1.96)(8.452)}{(2000)} \text{ lb/s.c.f.}$$

$$C_{.95} = 5.14 \times 10^{-7} \text{ lb/s.c.f.}$$

Therefore, the minimum detectable limit of Method 7 is 5.14×10^{-7} lb/s.c.f. If a laboratory obtains at a performance test result of 5.14×10^{-7} lb/s.c.f. using Method 7, there is only a 5% chance that the true nitrogen oxide emission was zero.

References

1. Youden, W. J., "The Collaborative Test", Journal of the AOAC, Vol. 46, No. 1, 1963, pp. 55-62.

2. Brownlee, K.A., Statistical Theory and Methodology in Science and Engineering, 2nd. Ed., Wiley, New York, 1965, pp 290-295.

3. Lindgren, B. W., Statistical Theory, MacMillan, New York, 1962, p. 89.

4. Lindgren, B.W., op. cit. p. 316.

5. Zeigler, R.K., "Estimators of Coefficient of Variation Using k Samples", Technometrics, Vol 15, No. 2, May, 1973, pp 409-414.

6. Mandel, John, "Repeatability and Reproducibility", Materials Research and Standards, ASTM Vol. 11, No. 8, August 1971, p. 12.

APPENDIX C. WALDEN THEORETICAL NO_x CONCENTRATION CALCULATION

The theoretical concentration of NO_x in the duct at the sample test section for the Cambridge test is given by

$$[C] = \frac{0.1183q}{Q_T + q} + K$$

where

$[C]$ = concentration of NO_x in the duct in lb/scf

q = NO flow from the gas doping system (acfm)

Q_T = theoretical volumetric flow in the duct (acfm)

K = 60×10^{-7} lb/scf NO_x due to combustion processes

The calculation of the flow due to the stoichiometric combustion of the No. 2 fuel oil is shown below.

$$Q_s = \left(\frac{0.871 \text{ moles of fuel}}{\text{min}}\right) \times \left(\frac{52.23 \text{ moles of combustion products generated}}{1 \text{ mole of fuel}}\right) \times \left(\frac{22.4\ell}{\text{mole}}\right) \times \left(\frac{1 \text{ ft}^3}{28.3\ell}\right) = 36.0 \text{ acfm from combustion only}$$

Calculation of the 52.23 moles of combustion products generated per mole of fuel is given in Tables C1 and C2.

Q_s is then corrected for the excess air present in the duct as follows:

$$Q_T = \frac{Q_s(100 - \%O_2 - \%CO_2)}{(100 - \%O_2 - \%CO_2) - (3.76)(\%O_2)}$$

where

Q_T = theoretical flow (acfm)

Q_s = flow from stoichiometric combustion (acfm)

$\%O_2$ = Fyrite reading

$\%CO_2$ = Fyrite reading

The values for $\%O_2$ and $\%CO_2$ over the course of the Cambridge test are shown in Table C3. The averaged Q_T result over a block of runs was used in computing the theoretical concentration.

The theoretical flows and calculated NO_x concentrations for each block of runs is shown in Table C4. A value of $K = 60 \times 10^{-7}$ lb/scf was used throughout as the background NO_x present due to the furnace combustion. A propagation of error analysis gave an error range of ±11% in the theoretical NO_x concentrations calculated by this procedure.

TABLE C1

OXYGEN CONSUMPTION FOR OIL COMBUSTION

Element	Wt. %†	Moles*	Moles* O_2 Needed
C	87.7	87.7/12	7.31
H	12.0	12.0/1	12.0/4 = 3.00
S	0.33	0.33/32	0.01
			TOTAL 10.32

*pounds moles/100 lb fuel.
†typical residual oil analysis.

TABLE C2

COMPONENT OF FLOW DUE TO STOICHIOMETRIC COMBUSTION

Species		Moles*
CO_2		7.31
H_2O	2.00 × 3.00 =	6.00
SO_2		0.01
O_2 (No excess air)		0
N_2†	10.32 × 3.77† =	38.91
	TOTAL	52.23

*pound moles/100 lb fuel.
†(nitrogen + argon)/oxygen ratio for dry air.

TABLE C3

WALDEN PILOT PLANT FIRING CONDITIONS

Date	Time	% O_2	% CO_2	Duct Temp. (°F) High	Duct Temp. (°F) Low	NO Doping Flow (ℓ/min)
12/11/72	12:20	8.5	9.0	480	275	2.25 ℓpm
	12:30	8.5	8.5	480	275	2.25
	1:40	8.5	8.5	475	290	2.25
12/12/72	11:30	9.5	9.5	480	275	2.1
	11:55	9.5	9.0	480	285	2.1
	12:15	9.0	9.5	480	285	2.1
	1:50	9.5	9.0	480	285	2.1
	2:10	9.5	9.0	480	285	1.75
	2:30	9.5	9.0	480	285	1.75
	3:15	9.0	9.0	475	290	1.75
12/13/72	11:30	9.0	8.5	475	275	1.11
	11:45	8.5	9.0	475	280	1.11
	11:50	9.0	9.0	475	285	1.11
	12:20	9.0	9.0	475	285	1.11
	1:00	9.0	9.0	475	285	1.11
12/14/72	12:30	9.0	8.5	480	-	0.52
	12:55	9.0	9.0	480	230	0.52
	1:00	9.0	9.0	480	240	0.52
	1:05	8.5	9.0	470	250	0.52
	1:20	9.5	9.0	470	260	0.52
	1:30	9.0	-	470	265	0.52
	1:45	8.5	8.0	470	275	0.52

TABLE C4

THEORETICAL CALCULATED NO_x CONCENTRATION RESULTS

Date	Theoretical Flow (acfm)	NO Doping Flow (acfm)	Calculated* Level of NO_x 10^{-7} lb/scf
12/11/72	58.70	0.0795	1660
12/12/72 a.m.	63.53	0.0742	1440
12/12/72 p.m.	63.16	0.0618	1216
12/13/72	60.74	0.0392	822
12/14/72	60.90	0.0184	417

*This value includes 60×10^{-7} lb/scf due to furnace combustion.